A Practicum in Behavioral Economics

A PRACTICUM IN BEHAVIORAL ECONOMICS

ARTHUR J. CAPLAN

Utah State University

Logan, Utah

DEDICATION

This book is dedicated to all the Homo sapiens I've come to know and love.

CONTENTS

PART I. SECTION 1 - THE IRRATIONAL QUIRKS OF *HOMO SAPIENS*

PART IV. SECTION 4 - EMPIRICAL STUDIES AND FIELD EXPERIMENTS

A PRACTICUM IN BEHAVIORAL ECONOMICS

ARTHUR J. CAPLAN

Media Attributions

- Technology connection concept

ACKNOWLEDGMENTS

The author acknowledges the following organizations and individuals for their support in helping to make this book a reality:

The Fulbright Specialist Program, the Bureau of Educational and Cultural Affairs, and the United States Department of State who provided me with the opportunity to teach a mini-course in behavioral economics to faculty at the Meiktila University of Economics in Myanmar during the summer of 2019. This experience provided the impetus and formed the basis of *A Practicum in Behavioral Economics*.

My students and administrators at Meiktila University of Economics, who intrepidly crossed the language barrier and willingly participated in the book's maiden voyage. They showed me just how powerful the approach used in *A Practicum in Behavioral Economics* can be in bringing the field of behavioral economics to life.

Nick Gittins and Stephanie Western, Open Educational Resources Coordinator and Program Manager, respectively, and their colleagues and staff at the University Libraries Digital Initiatives, Utah State University, who helped turn my initial draft of the book into a reality.

And last, but certainly not least, my wife, Frances Caplan, who encouraged me to make the trip to Myanmar and to take a chance on transforming the experience into the book you now hold.

THIS BOOK'S APPROACH

This book's approach is premised on a simple assumption: because behavioral economics is foremost a "test-and-learn" field of scientific inquiry that evolves according to experimental outcomes and practical, policy-orientated applications of the knowledge garnered from these outcomes, so too should students test-and-learn. Studying and practicing behavioral economics should occur simultaneously, which, in turn, suggests a course taught more according to a practicum approach than in a traditionally styled lecture format. As such, the book's information and lessons are presented in a succinct and precise format.

The goal of this textbook is to help students experience behavioral economics through actual participation in the same experiments and economic games that have served as the foundations for, and shaped the contours of, the field. With the help of this book, students have the opportunity to learn behavioral economics firsthand and, in the process, create their own data and experiences. They will learn about themselves—about how they make private and public choices under experimental conditions—at the same time as they learn about the field of behavioral economics itself. They will be both the subjects and students of behavioral economics. What better way to learn?

HOMO ECONOMICUS VS. *HOMO SAPIENS*

For ease of reference and exposition, we henceforth refer to the type of individual construed by the traditional rational-choice model as *Homo economicus*, a peculiar subspecies of human beings that is unfailingly omniscient, dispassionate, and self-interested when it comes to making choices. *Homo sapiens*, on the other hand, represents the rest of us—the often-flawed reasoners and sometimes-altruistic competitors who are prone to making decisions based primarily on emotion and heuristics.[1],[2]

THE TEXTBOOK'S DIFFERENT SECTIONS

The textbook consists of four sections that, taken together, portray in full the eclectic methodologies comprising the field of behavioral economics. Sections 1 and 2 present the thought and actual

1. *Homo economicus* is Latin for "economic man." Persky (1995) traces its use back to the late 1800s when it was used by critics of John Stuart Mill's work on political economy. In contrast (and, as we will see, with no small touch of irony) *Homo sapiens* is Latin for "wise man." For a deep dive into evolution of *Homo sapiens*, particularly from the start of the Cognitive Revolution 70,000 years ago, see Harari (2015).

2. We have all heard the saying that "words matter." The titles and descriptions we use to distinguish people and their behaviors (e.g., *Homo economicus* vs. *Homo sapiens*) can reinforce or diminish behaviors such as pride in cultural heritage, respect for the living world, and trust in community, a process known as "crowding out" of "intrinsic motivation and commitment." As an example of this phenomenon, Bauer et al. (2012) asked participants in an online survey to imagine themselves as one of four households facing a water shortage due to a drought affecting their shared well. The survey assigned the label "consumers" to half of the participants and "individuals" to the other half. Those imagining themselves as consumers reported feeling less personal responsibility to reduce their water demand, and less trust in others to do the same, than did those referred to as individuals. As we are about to learn, behavioral economics is all about exposing these types of "framing effects" existing in the "real world" inhabited by *Homo sapiens*.

laboratory experiments that have formed key pillars of the field, such as those experiments depicted in Examples 1 and 2 in the book's Introduction section. The thought experiments in Section 1 are, for the most part, re-castings of the simple cognitive tests devised by psychologists and economists over the past three-to-four decades to illustrate the fallacies, miscalculations, and biases distinguishing *Homo sapiens* from *Homo economicus*. Similarly, the laboratory experiments presented in Section 2 are, for the most part, re-castings of the seminal experiments conducted by Kahneman and Tversky (among many others). These experiments helped motivate the revised theories of human choice behavior, such as Kahneman and Tversky's (1979) Prospect Theory, which form another pillar of behavioral economics. Alongside these experiments, Section 2 presents the revised theories of human choice behavior with varying degrees of rigor. This is where the theoretical bases of *Homo economicus*' rational choice behavior are examined, and where key refinements to this theory are developed—theoretical refinements underpinning the myriad departures from rational choice behavior we witness *Homo sapiens* make in this section's laboratory and field experiments (and which are examined further in Sections 3 and 4).

Section 3 submerses the student in the world of behavioral game theory. Here we explore games such as Ultimatum Bargaining presented in Example 5. We follow Camerer (2003)'s lead, first by characterizing the games analytically (i.e., identifying solution, or equilibrium, concepts that are predicted to result when members of *Homo economicus* play the games), and then by discussing empirical results obtained from corresponding field experiments conducted with *Homo sapiens*. It is within the context of these games and field experiments that theories of social interaction are tested concerning *inter alia* trust and trustworthiness, honesty, fairness, reciprocity, etc. As with the thought and laboratory experiments presented in Sections 1 and 2, the games and field experiments presented in Section 3 are meant to be replicated with students as subjects and the instructor as the experimenter, or researcher.

Finally, Section 4 wades into the vast sea of empirical research and choice architecture. Here the student explores studies reporting on (1) the outcomes of actual policy nudges, such as the SMarT retirement-savings plan presented in Example 3 of the Introduction, (2) analyses of secondary datasets to test for choice behavior consistent with the revised theories discussed in Section 2, such as the test for loss aversion in Example 4 of the Introduction, and (3) analyses of primary datasets obtained from novel field experiments to further test the revised theories. The main purpose of this section is not only to introduce the student to interesting empirical studies and policy adaptations in the field of behavioral economics, but also, in the process, to incubate in the student an abiding appreciation for the obscure settings that sometimes lend themselves to such study.[3]

THE TEXTBOOK'S DIFFERENT LEVELS OF RIGOR

Because the mathematical and computational rigor of material presented in this textbook varies throughout, particularly in Sections 2 – 4, the extent of the rigor used in the presentation of a given topic is indicated with superscripts. Topics without a superscript are considered basic and universal enough that backgrounds in economics, mathematics, or statistics are not required for the reader to understand the material. Topics with a single asterisk (*) indicate that higher mathematical reasoning skills are recommended for the reader to fully grasp the material. Topics with a double

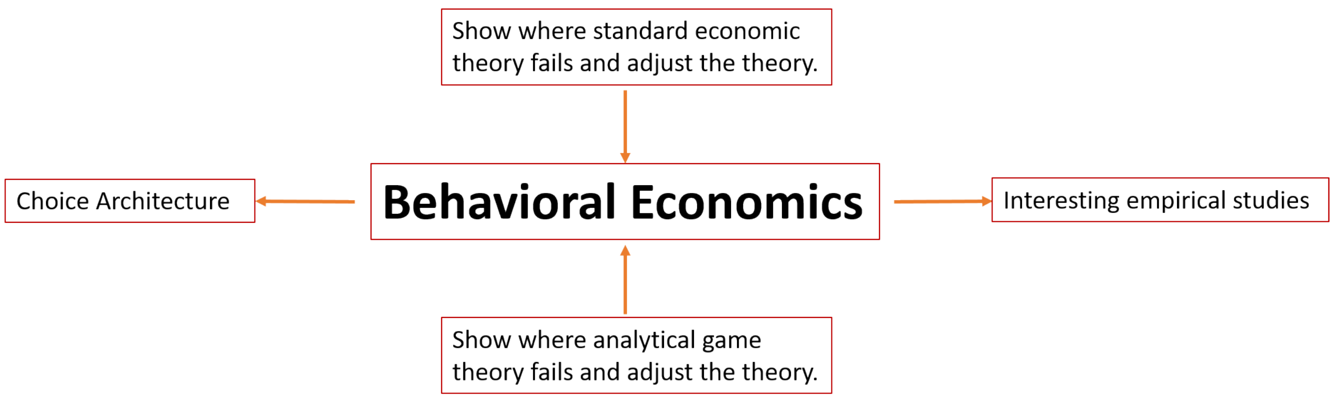
3. Our approach to studying behavioral economics is focused on the underlying laboratory experimentation and behavioral games that form the bedrock of the field. As such, we eschew delving into related fields such as neuroeconomics and auction theory. See Cartwright (2018) and Just (2013) for introductions to the former and latter fields, respectively.

asterisk (**) indicate that either higher economic or statistical reasoning skills, whichever the case may be, are recommended. And lastly, topics with the dreaded triple asterisk (***) indicate that both higher economic/statistical and mathematical computational skills are likely required to fully grasp the material. Both students and instructors should bear these indicators in mind.

For example, none of the topics presented in Section 1 are superscripted, implying that students from varied academic backgrounds should be able to fully understand the material presented. Single asterisks (*) first appear in Section 2, Chapter 3, indicating that the discussions of the Principle and Additional Rationality Axioms pertaining to *Homo economicus* will likely be more easily comprehended by students with higher mathematical reasoning skills. The double asterisk appears later in the same chapter when the topic of *Homo economicus* and the expected utility form is presented, and the bug-a-boo triple asterisk first appears at the end of Chapter 3, demarcating the topic of intertemporal choice.

THINKING DIAGRAMMATICALLY

For those who prefer thinking diagrammatically, the figure below illustrates how these four sections relate to, and help define, what we thus far understand to be the field of behavioral economics.



The two boxes with arrows pointing inward toward **Behavioral Economics** can be thought of as the "inputs" to our understanding of the field. The box enclosing the statement, "Show where standard economic theory fails…." represents Section 1 of the guidebook, and "….adjust the theory" pertains to Section 2. The box enclosing the statement, "Show where analytical game theory fails and adjust the theory" represents Section 3. In contrast, the two boxes with arrows pointing outward from **Behavioral Economics** can be thought of as "outputs" in the sense of Choice Architecture (e.g., the SMarT retirement-savings plan described in Example 3 of the Introduction) and "interesting empirical studies" (e.g., the PGA study described in Example 4 of the Introduction). These two areas of interest are explored in Section 4.

THE TEXTBOOK'S APPENDICES

Appendix A at the end of the book includes example Response Cards for the experiments and games presented in Sections 1-3. I am old-fashioned when it comes to collecting student responses—I print out a response card for each student for each experiment or game, have the students fill in their responses, and then pass around a "collection box" for each student to place his or her card in.

Student ID numbers on the response cards could be their names or their university ID numbers. Or, if you wish to align the students' responses with more demographic information obtained from a survey instrument administered on the first day of class, you might consider randomly assigning the students individual course identification (CID) numbers on the first day of class. The CID numbers would not be tied to the students' names or their university identification numbers. In this case, in order to preserve their anonymity, you would be precluded from basing the students' course grades upon their performance in the experiments or their responses to the survey instrument. Appendix B includes a copy of a socio-demographic survey that could be administered to students the first day of class, after having randomly assigned their CID numbers. This information would be useful when it comes to analyzing the data obtained from the experiments and games. Again, because the surveys are linked to CID numbers rather than student names or university identification numbers, student anonymity regarding the survey instrument is ensured.

Appendix C includes examples of presentation slides used for lectures and as guides for the experiments and games as students proceed to participate in them. Appendix D includes examples of course outlines designed for courses targeting economics and non-economics majors, respectively. And a Linkages Matrix is provided in Appendix E. This matrix provides a structure for identifying connections between the various concepts presented in Chapters 1 – 4 and the experiments, games, and empirical studies discussed in Chapter 6 and later in Section 4.

Media Attributions

THE BOOK'S GENESIS AND TIPS ON HOW TO USE IT

This textbook evolved during the summer of 2019. In February of that year, I accepted a Fulbright Specialist position to teach an intensive short-course in behavioral economics to the faculty at Meiktila University of Economics in the southeast Asian nation of Myanmar. In the project description, Dr. Thida Kyu (PhD economist and Pro-Rector of the university), explained that because it had been cut off from the Western world for so long (roughly 50 years) and had only recently attained (nominal) civilian control of the nation's government in 2016, Myanmar had a lot of catching up to do, particularly regarding the functioning of its academic institutions. Dr. Kyu was aware of this new field called behavioral economics. She believed its lessons would not only enlighten her faculty and their students, but might also help nudge her country's fight against poverty onto a more enlightened path policy-wise. I took this to mean, rightly or wrongly, that Dr. Kyu was not looking for another lecture-orientated course, a mere overview of the history, methodologies, and findings of behavioral economics. Rather, her faculty needed a practitioner's guide, a course that would, as much as possible, engage them with the field's methodologies and findings through actual practice and firsthand experience—a course that would get them in on the proverbial ground floor of this relatively new field of inquiry.

I began my preparations for the course by doing what I always do when assigned to teach a new course. I sought out existing textbooks. Over the course of my career, I've been fortunate to have a wide variety of textbook selections for the fields of environmental and resource economics and microeconomic theory. But not this time. It became apparent almost immediately that if I were to prepare a course geared more toward the practice of behavioral economics, I would need to cobble together material from a host of disparate sources. The book you now hold is the result of this 'cobbling' process. It melds Kahneman's, Tversky's, and Thaler's seminal works (along with several other key theoretical and experimental advancements published in a wide variety of journals over the past 50-plus years) with Camerer's (2003) behavioral game theory text and William Spaniel's (2011) introductory textbook on analytical game theory. The book also draws from Kahneman's (2011), Ariely's (2008), and Thaler and Sunstein's (2009) *New York Times* bestsellers *Thinking, Fast and Slow*, *Predictably Irrational*, and *Nudge*, respectively, and to lesser extents, from Levitt and Dubner's (2005), Gladwell's (2002), and Harford's bestsellers *Freakonomics, The Tipping Point*, and *Messy*.

In addition to the value-added that comes from having incorporated these works into a single text—in some cases, rendering explicit representations of experiments the authors have merely mentioned and in other cases drawing directly from the original sources cited by the authors—I have included material from works that I consider to be worthy representations of the breadth of behavioral economics as a field of inquiry. In the end, we have before us a book that guides the student through this field no differently than a well-researched guidebook helps the intrepid international traveler navigate a foreign country's main attractions, and helps the traveler gain knowledge of (and hopefully appreciation for) the country's history and cultural uniqueness.

As such, this book is not necessarily meant to be read by students from cover-to-cover in chronological order (i.e., first covering the material in Section 1, then the material in Section 2, and so on). Rather, it is possible that what works best for your students is for them to be introduced to the material in a piecemeal fashion. For example, when I taught the course in Myanmar, I included in each three-hour lecture an experiment or two from Section 1 coupled with some of the economic theory presented in Section 2, and either a game from Section 3 or a discussion of empirical research or choice architecture from Section 4. This helped the students engage with each of these facets of behavioral economics for the duration of the course. It also precluded me from front-loading the often fun-filled experiments and games, and leaving Section 2 and 4's more lecture-orientated (dare I say less-entertaining?) discussions of the theory and empirical research and choice architecture for the last few weeks of the course.

Perhaps most importantly, drawing from more than one section of the book in each lecture facilitates the connecting of an outcome from a Section 1 thought experiment to a laboratory experiment (and an associated, revised economic theory) in Section 2, or connecting an outcome from a Section 3 game to a corresponding result from a Section 4 empirical study. Indeed, the gamut of potential connections that can be made across the topics presented in the different sections of the book is almost limitless. Since economists tend to deal better with finiteness than infiniteness, Appendix E provides what I call a "linkages matrix," which, provides a structure for identifying connections between the various concepts presented in Chapters 1 – 4 and the experiments, games, and empirical studies discussed in Chapter 6 and Section 4. This matrix is meant to serve as an aid for instructors who adopt this type of piecemeal approach to teaching the course.

For example, one of the thought experiments presented in Section 1 exemplifies what Kahneman (2011) and Kahneman and Tversky (1984) originally labeled a "framing effect," which in turn can lead to a host of biases in choice behavior, such as confirmation bias and representative bias.[1] Accordingly, in Section 3 we could discuss results from a field experiment that shows how framing the Ultimatum Bargaining game as a "seller-buyer exchange" encourages self-interest (i.e., behavior expected from *Homo economicus*), while framing the game as a common-pool resource encourages *Homo sapiens*-like generosity. Empirical research presented in Section 4 demonstrating "loss aversion" on the part of public school teachers in Chicago can also be considered an example of a framing effect as first introduced in Section 1 since the timing (i.e., framing) of bonus payments made to teachers based on improved student performance is the mechanism eliciting the loss-averse behavior. Linkages like these abound across the four sections.

If instead of adopting the piecemeal approach to teaching the course, the instructor prefers a more traditional, chronological approach to presenting the material as laid out in Sections 1 – 4, the annotated course outline provided in Appendix D offer guidance. One outline is designed for a course targeting economics majors, the other for a course targeting non-majors. The main difference between the two outlines is that the former allocates more time to the economic concepts and theories presented in Chapters 3 and 4, while the latter emphasizes the material covering human quirks (e.g., heuristics, biases, and effects presented in Chapters 1 and 2). As the course outlines for both types of courses indicate, the instructor chooses the specific effects, biases, theoretical material, experiments, games, and empirical studies that will be covered in lectures.

In concert with the course outline, figuring out how best to grade students in a course like this can be a challenge, particularly if you decide to administer a demographic survey (Appendix B) on the first day of class. In this case, preserving student anonymity becomes an issue. To deal with this issue, consider creating two separate spreadsheets for the course. One spreadsheet compiles the students'

1. Indeed, several of the other effects presented in Section 1, e.g., anchoring effect and halo effect, can be thought of as special cases of a framing effect.

survey responses and outcomes from the experiments and games. This spreadsheet is linked to the students' randomly assigned course ID (CID) numbers. The other spreadsheet, which is linked to their university student ID numbers and their names, compiles their performances on quizzes, homework, and exams assigned throughout the semester.

At the risk of sounding draconian, this is a course where it may make sense to base upwards of 50% of a student's grade upon their in-person attendance, which would entail carefully taking role at the beginning of each class. If the class meets 30 times face-to-face during the semester, for example, their grade attributable to attendance would then drop by 3.33 percentage points for each missed class (excused absences withstanding). Granted, students who foresee having difficulty attending class in-person throughout the semester would likely choose to drop the course immediately. For those students who remain, the remaining 50% of their course grade would then be based upon their quizzes, homework, and exam scores.

The issue of how best to convey written information to the student a priori (i.e., before conducting a given experiment or game) also looms large in a participatory-learning setting such as this, especially if the instructor desires to obtain unbiased responses from the students (or more practically, to control for potential biases). For example, the first set of thought experiments presented in Section 1 is meant to demonstrate firsthand to the students the extent to which automatic, knee-jerk responses from what Kahneman (2011) identifies as the System 1 portion of the brain can result in miscalculations. Students who choose to read ahead (small in number though these types of students may be) potentially skew the distribution of responses away from its otherwise true representation of these miscalculations. Such skewness may be tolerable for strictly educational purposes, where the goal is to demonstrate that at least a certain percentage of students are prone to miscalculation. But if the instructor also hopes to compile student responses into a dataset amenable for statistical analysis, then this type of potential bias draws into question the validity of the data.[2]

To help control for potential biases associated with students having read ahead about the game or experiment they are now participating in, I recommend including the following question on each Response Card: "Did you read about this topic ahead of time?" (see Appendix A). Answers to this question provide a control for the level of student foreknowledge, which is the potential bias of concern.

I am personally unaware of any studies that have looked at how well students learn the lessons of behavioral economics in a cumulative sense over a span of time (e.g., an entire semester) and across a variety of experiments and games. In other words, I know of no studies that estimate the extent to which individuals who begin a course in behavioral economics as bona fide *Homo sapiens* evolve toward "*Homo economism*" in their individual and social choices. The pedagogy promoted in this textbook—in particular, the data it generates—offers instructors the opportunity to empirically test the hypothesis that students make this evolution.

2. Note that this potential biasedness problem also extends to the laboratory experiments of Section 2 and games of Section 3.

A (VERY) BRIEF HISTORY OF THE ORIGIN OF BEHAVIORAL ECONOMICS

Most historical accounts trace the origin of behavioral economics as far back as Adam Smith's *The Theory of Moral Sentiments*, published in 1759 (Loewenstein, 1999; Camerer and Loewenstein, 2004; Angner and Loewenstein, 2012; Thaler, 2016).[1],[2] As Camerer and Loewenstein (2004) point out, Smith was the first to propose that we humans derive more disutility (i.e., unhappiness) from losses than we do utility (happiness) from gains, a conjecture of "loss aversion" that later formed the basis of Kahneman and Tversky's (1979) Prospect Theory. And so, in the mid-18th century, just as economics began to be considered a separate discipline, it appeared as though economic thought would necessarily evolve in tandem with our understanding of human psychology. However, by the turn of the 20th century and the onset of the neoclassical revolution, economists began turning away from what was considered to be the inherently unscientific nature of psychological analysis, ultimately leading to the positivistic theories of human choice behavior posited by the likes of Veblen, Hicks, Stigler, Menger, Jevons, and Walras (to name but a few), and later the normative and descriptive models of expected and discounted utility proposed by post-war neoclassicists von Neumann, Morgenstern, and Samuelson.[3]

Because of the strong assumptions underpinning the expected utility and discounted utility models (e.g., the Independence Axiom and exponential discounting, respectively), critics such as Allais, Ellsberg, Markowitz, and Strotz had, by the middle of the 20th century, identifed anomalous implications associated with these models. These implications would later be demonstrated in the famous laboratory experiments of Kahneman, Tversky, and Thaler (Camerer and Loewenstein, 2004). At around the same time as Kahneman and Tversky were running their experiments, developments in the field of cognitive psychology—known as "behavioral decision research"—suggested promising new directions for explaining choice behavior as a consequence of the brain's information-processing

1. *The Theory of Moral Sentiments* was Smith's lesser-known book. He is best known for *The Wealth of Nations*, published roughly 15 years later in 1776, where he coined the now famous term "invisible hand." Other important works commenting on the psychological underpinnings and determinants of utility—the bedrock concept of early 20th century neoclassical economics—include Bentham's *An Introduction to the Principles of Morals and Legislation* (1789) and Edgeworth's *Theory of Mathematical Psychics* (1881) (Camerer and Loewenstein, 2004).
2. Heukelom (2006) traces the origin of behavioral economics back further to the gambling problems proposed by French nobleman-gambler Chevalier de Méré in 1654. Perhaps the most famous gambling problem, the St. Petersburg paradox, was coined by Daniel Bernoulli in 1738 in his *Commentaries of the Imperial Academy of Science of Saint Petersburg*. Bernoulli's solution to this gambling problem—the maximization of expected utility—rested upon the assumption of diminishing marginal utility of wealth (Heukelom, 2006).
3. As Camerer and Loewenstein (2004) point out, economists such as Irving Fisher and Vilfredo Pareto still stressed the role of psychology in choice behavior in the early part of the 20th century. In the latter part of the century, economists George Katona, Harvey Leibenstein, Tibor Scitovsky, and Herbert Simon—fathers of what is affectionately known as "old behavioral economics"—similarly stressed the role of psychology and bounded rationality as constraints on choices (Angner and Loewenstein, 2012). As Heukelom (2006) points out, economics and psychology ultimately go separate ways, the former employing Friedman's positive-normative distinction, the latter using Savage's normative-descriptive distinction.

capability.[4] As described in Angner and Loewenstein (2012), this parallelism between advancements in cognitive psychology and economic experimentation, along with the fact that cognitive science as a separate field of inquiry arose in opposition to the field of behavioralism in psychology, suggests that the label "behavioral economics" is arguably a misnomer. Perhaps it would be more accurate to dub the field "cognitive economics."[5]

4. See Hastie and Dawes (2001) for a nice discussion of behavioral decision research.

5. Kahneman (2011) provides an accessible account of how our brain's information-processing capability drives the misconceptions and miscalculations that ultimately lead to the fallible heuristics and biases that he, Tversky, and Thaler (among others) have both documented in their experiments and subsequently used as grist for their alternative theories of choice behavior. These theories, explored in Section 1 of this book, are in turn the mainstay of behavioral economics.

INTRODUCTION

The advancement of science in large part depends upon observation of behavior that has either never been encountered before or, if previously encountered, remains inadequately explained. Observation of what is presently inscrutable propels scientific inquiry. In the fields of cosmology and physics, for instance, observations of particles and the "arrow of time" have propelled the search for our universe's origin and evolution (Hawking, 2017). In the field of neuroscience, observation of the genetic barcode of a mouse's brain cells has enlightened our understanding of how human cells mature with age, how tissues regenerate, and how disease impacts these processes (Pennisi, 2018). And so it is with what has come to be known as behavioral economics, a field of inquiry melding psychology's long-running exploration of human cognition and social norms with the long-standing axioms of omniscient rationality that economists have traditionally ascribed to human choice behavior. Behavioral economics is the long-awaited advancement in economic theory and experimentation that involves both deconstructing and reconstructing the economist's rational-choice, neoclassical model to better explain the choices individuals actually make on a daily basis, and ultimately to better inform public policy. Through their keen observations of human choice behavior in a wide variety of contexts, behavioral economists have propelled scientific inquiry.

As aptly pointed out by Samson (2019), observations of choice behavior in both private and social settings demonstrate the extent to which human decisions are influenced by context, including how choices are presented to us. The observations demonstrate ways in which our choice behavior is subject to cognitive biases, emotions, heuristics, and social influences. Because these biases, emotions, and influences have, in turn, been shown in a myriad of well-designed laboratory and field experiments and empirical studies to govern choice behavior in ways unpredicted by economists' rational-choice models, we cannot help but celebrate the emergence of behavioral economics as a separate field of inquiry. In some sense, behavioral economics can be thought of as an overt partnership between the complementary fields of psychology and economics—a natural blending of the former's insights on human cognition and the latter's focus on choice behavior. As we will learn in this textbook, behavioral economics is a beacon, not only for the revision and generalization of key features of the economist's rational-choice model of human behavior but also for what Thaler and Sunstein (2009) have popularized as "nudges" that can improve the outcomes of public policymaking.

Five examples depict the reach of behavioral economics as a separate field of inquiry and illustrate its emergence as a canon of human choice behavior. The first two examples demonstrate precisely how this behavior deviates from the economist's rational-choice model in the confines of laboratory and field experimentation. The third example demonstrates how policymakers have leveraged these experimental findings to nudge private decisions toward more preferable social outcomes. The fourth example shows how researchers have tested the findings with real-world data obtained from unexpected places. And the fifth example demonstrates what is known as "behavioral game theory," outcomes of well-known economic games that depart from theoretical predictions, sometimes in rather dramatic fashion.

EXAMPLE 1

The Invariance Axiom is central to expected utility theory, i.e., rational choice behavior under uncertainty. Simply put, the axiom holds that an individual's preference ordering of different lotteries (e.g., ranking from most to least preferred lottery) does not depend upon (i.e., is invariant to) how the lotteries are described to the individual. Kahneman and Tversky (1984) test this axiom with a simple experiment involving two subject groups, each group totaling roughly 150 students.

Group 1 was presented with the following lottery:

> Imagine that your hometown is preparing for the outbreak of an unusual disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
>
> If Program **A** is adopted, 200 people will be saved.
>
> If Program **B** is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.
>
> Which of the two programs would you favor?

Group 2's lottery was this:

> Imagine that your hometown is preparing for the outbreak of an unusual disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
>
> If Program **C** is adopted, 400 people will die.
>
> If Program **D** is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.
>
> Which of the two programs would you favor?

If you look closely at the two lotteries, you will note that they are identical. Program **A** from Group 1's lottery is identical to Program **C** from Group 2's lottery, and Group 1's Program **B** is identical to Group 2's Program **D**.[1] Thus, we expect the percentages of Group 1 students choosing among Programs **A** and **B** in their lottery to be roughly equal to the corresponding percentages of Group 2 students choosing among Programs **C** and **D** in their lottery. This would be in keeping with the Invariance Axiom.

Instead, Kahneman and Tversky (1984) found that 72% of Group 1 students chose Program **A** and

1. The latter identity results because a one-third probability that 600 people will "be saved" under Program **B** means 0.33 x 600 = 200 people are expected to be saved, which is the same number of people who are not expected "to die" under Program **D**. Similarly, the 0.67 x 600 = 400 people who are not expected to be saved under Program **B** is the same number of people who are expected to die under Program **D**. Hence the two lotteries are indeed identical.

28% chose Program **B**, while only 22% of Group 2 students chose Program **C** and 78% chose Program **D**, a dramatic refutation of the Invariance Axiom. The authors concluded that because the "reference points" of the two lotteries differed in this experiment—Group 1's is that people are "saved" and Group 2's is that people "die"—the Invariance Axiom was not necessarily destined to hold in this context, which runs counter to the rational-choice model's presumption that the axiom holds in any context. As we will see, this insight led to Kahneman and Tversky's notions of "reference dependence" and "framing" in human choice behavior; notions which had been ignored by the rational-choice model, yet are crucial to our understanding of how humans make decisions under uncertainty. In short, context matters.

EXAMPLE 2

Heath and Tversky (1991) engaged roughly 200 subjects in the following lottery:

> Choose between lotteries **A** and **B**:
>
> **A** A stock is selected at random from the New York Stock Exchange. You guess whether its price will go up or down at close tomorrow. If your guess is correct you win $100.
>
> **B** A stock is selected at random from the New York Stock Exchange. You guess whether its price went up or down at close yesterday. You cannot check the newspaper or online. If your guess is correct you win $100.

Bearing in mind that the internet was not yet in widespread use in 1991, and thus lottery **B** was indeed failsafe, we would expect the subjects to be indifferent between the two lotteries, resulting in a 50-50 split of those choosing **A** versus **B**.[2] Instead, 67% of the subjects chose lottery **A** and 33% percent chose **B**, which supports what the authors labeled a "competency effect." The supermajority of subjects preferred the future bet because their "relative ignorance" was easier to defend this way. In a sense, they appeared less incompetent by choosing lottery **A**.

EXAMPLE 3

This example highlights a nudge to public policy (in the form of a single company's benefits policy) that leverages our understanding of framing from Example 1. In particular, the example explores how framing a new retirement-savings program appropriately can overcome what is known as "status quo bias" among a company's employees.

As Thaler and Benartzi (2004) point out, US companies have been switching their retirement plans over time from *defined-benefit* to *defined-contribution plans*. Under defined-contribution plans, employees bear more responsibility for making decisions about how much of their salaries to save. Employees who participate in a given plan at a very low level save at less-than-predicted life-cycle (i.e., rational) savings rates. One explanation for this irrational behavior is a lack of self-control among low-saving employees, suggesting that at least some of these workers are making a mistake and would welcome help in making decisions about their retirement savings. It could also be that some

2. The technical terminology for the rational-choice axiom, in this case, is "additivity of subjective probability."

employees suffer from the competency effect portrayed in Example 2. Either way, employees tend to exhibit status quo bias when it comes to optimizing their retirement-savings plans.

To counteract this problem, Thaler and Benartzi (2004) devised a new savings plan for a large company called the Save More Tomorrow (SMarT) plan. The essence of the plan is straightforward: people commit now to increasing their savings rate later (i.e., each time they get a pay raise). As will be explained further in Section 4 of this textbook, the authors found that the average saving rates for SMarT participants increased from 3.5% to 13.6% over the course of the plan's first 40 months, while employees who chose an alternative retirement plan increased their saving rate to a lesser extent. Those who declined both the SMarT and alternative plans saw no increase in their savings rates.

The question naturally arose as to how the company might entice more of its employees to enroll in the SMarT plan. One suggestion was to frame the choice of retirement plans as an "opt-out" rather than an "opt-in" decision. Under opt-out, new employees are automatically enrolled in the SMarT plan and therefore must take it upon themselves to switch to another plan. Opt-out ingeniously harnesses employees' natural tendencies toward status quo bias for their own betterment (at least regarding retirement savings decisions).[3]

EXAMPLE 4

Pope and Schweitzer (2011) explore whether reference dependence (such as that described in Example 1), and "loss aversion," (which is one of behavioral economics' most renowned discoveries in laboratory experiments), are present in the behavior of professional golfers.[4] Loss aversion governs choice behavior when an individual perceives the pain of losing as more powerful than the pleasure of winning (or, gaining). Loss-averse individuals are more willing to take risks or behave dishonestly to avoid a loss than to achieve a gain (behavioraleconomics.com, 2019).

As Pope and Schweitzer (2011) point out, golf provides a natural setting to test for loss aversion because golfers are rewarded for the total number of strokes they take during a tournament, yet each hole has a salient reference point, putting for par. Loss-averse golfers suffer more psychologically from scoring "over par" (bogeying) on any given hole than "under par" (birdying). The authors analyzed over 2.5 million putts measured by laser technology and found evidence that even the best golfers—including Tiger Woods in his heyday—show evidence of loss aversion. Specifically, when PGA golfers are under par on any given hole (i.e., putting for a birdie), they are 2% less likely to make the putt than when they are putting for par or are over par (i.e., putting for a bogey).

EXAMPLE 5

The Ultimatum Bargaining game is one of the most widely tested games in the history of behavioral game theory. It has been tested with students in the US and Europe, as well as tribes in Africa, the Amazon, Papua New Guinea, Indonesia, and Mongolia. The game is described as follows:

3. The opt-out approach has been shown to work in other instances as well, most famously for organ donor programs. Davidai et al. (2012) point out that Spain, Belgium, Austria, and France have among the highest organ-donation consent rates worldwide, precisely because they use opt-out defaults (known as "presume consent") when it comes to registering citizens in their respective programs. To *not* donate their organs upon death, citizens must take it upon themselves to opt out (i.e., they must overcome status quo bias with respect to donating their organs).

4. If you are wondering why professional golfers, it is because of the plethora of data that exists from the various Professional Golfers Association (PGA) tournaments held each year.

Two players – a Proposer and a Responder – bargain over \$10. The Proposer offers some portion, x, of the \$10 to the Responder, leaving the Proposer with \$(10-x). If the Responder accepts the offer, then she gets \$x and the Proposer gets \$(10-x). If the Responder rejects the offer, both players get nothing.

Camerer (2003) points out that by going first, and because the game is played in "one shot," the Proposer has all of the bargaining power. Therefore, we should expect, per the rational-choice model, that the Proposer will exploit the fact that a similarly self-interested Responder will take whatever is offered. The Proposer should thus offer an \$*x* very close to \$0.

Instead, in a multitude of experiments conducted worldwide, Proposers typically offer roughly half of the total. Offers of roughly 20% are rejected about half of the time as punishment for what Responders interpret as Proposers not having behaved fairly. Variants of the game have considered more than one Proposer, repeated play between a Proposer and Respondent with "stranger matching" (i.e., new pairings among the pool of subjects), higher stakes, and added risk associated with the Responder not knowing for certain what the stakes are. Again and again, the behavior of participants in the game deviates from the expected, rational outcome.

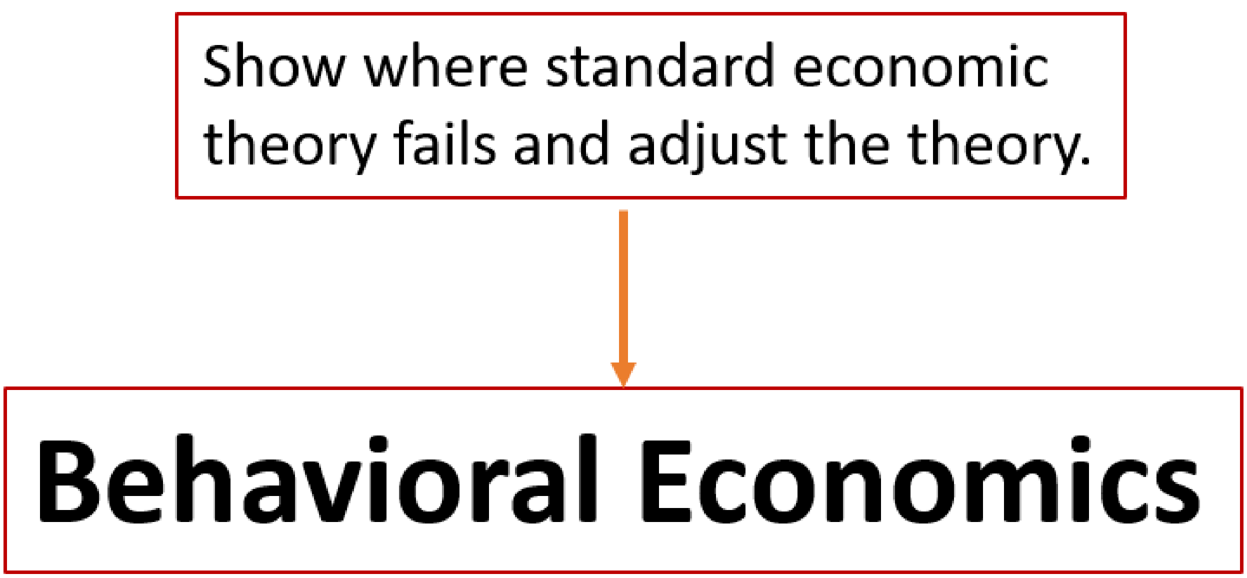
To reiterate and sum up our introductory remarks, human beings do not always behave as the self-interested, net benefit maximizing individuals with stable preferences that the traditional rational-choice model of economic decision making would have us believe. Let's face it. Most of our choices are not the result of careful deliberation. We are influenced by readily available information in our memories and automatically generated, salient information in the environment. We live in the moment and thus tend to resist change, are poor predictors of future behavior, subject to distorted memory, and affected by physiological and emotional states of mind. We are social animals with social preferences, susceptible to social norms and a need for self-consistency (Samson, 2019). All of this we sense intuitively; these are normal human behaviors. Behavioral economics studies how this normality plays out in economic and social contexts, and in the process identifies where traditional rational-choice theory has fallen short of correctly predicting individual and social choice behavior.

PART I.

SECTION 1 - THE IRRATIONAL QUIRKS OF *HOMO SAPIENS*

In the Introduction, we alluded to the fact that behavioral economics as a separate field of inquiry serves three main purposes. First, it responds to the limitations of the neoclassical paradigm of *Homo economicus*, an idealized version of a human being who behaves eminently rationally: knowingly and selfishly, with unlimited computational capacity, never making systematic mistakes. Second, behavioral economics provides a clearer understanding of how *Homo sapiens* actually behave, given all of our irrational quirks—our miscalculations, misjudgments, inconsistencies, contradictions, illusions, moods, biases, fallacies, and so on (if that isn't enough).[1] Third, the field of behavioral economics proposes adjustments to the theories historically predicated on the choice behaviors scripted for *Homo economicus*.[2]

Referring to the diagram presented in the This Book's Approach section, Section 1 of the textbook pertains to the diagram's upper portion.



1. Ariely (2008) lumps all of these irrational quirks into what he calls predictable irrationality.
2. The revisions are decisive enough as to be considered stand-alone theories themselves (e.g., Kahneman and Tversky's (1979) Prospect Theory), which will unfold later in this section.

Here, we demonstrate how standard economic theory fails by highlighting the major disconnects between behavior predicted of *Homo economicus* and that displayed by *Homo sapiens* (i.e., you, me, and your fellow students). We begin by evincing our proclivities for committing the miscalculations, misjudgments, etc. mentioned above through thought and laboratory experiments that you will participate in, as well as through brief discussions of outcomes from interesting laboratory and field experiments published in academic journals. Section 4 includes more in-depth discussions of some of the studies mentioned here.

Periodically, we will delve into the standard economic theory being tested by the laboratory experiments, and we will pinpoint how this theory has been revised in light of the outcomes of these experiments. This material is perhaps best described as the gist of behavioral economics. Unless otherwise indicated, the material in this section has either been taken directly from Kahneman (2011) or is based on discussions therein.

Media Attributions

CHAPTER 1.

MISCALCULATIONS, COGNITIVE ILLUSIONS, MISJUDGMENTS, AND 'EFFECTS'

We begin by considering some well-known miscalculations that bedevil and typify *Homo sapiens*.

(RELATIVELY SIMPLE) MISCALCULATIONS

A baseball bat and ball together cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?

If it takes 5 machines 5 minutes to make 5 football helmets, how long would it take 100 machines to make 100 football helmets — 100 minutes or 5 minutes?

All daises are flowers. Some flowers fade quickly. Thus, some daises fade quickly. Is this syllogism valid?

In a lake, there is a patch of waterlilies. Every day the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take the patch to cover half the lake — 24 days or 47 days?

Answers

Box 1 — The ball costs $0.05.

Box 2 — It would take 5 minutes.
Box 3 — The syllogism is not valid.
Box 4 — It would take 47 days.

Homo economicus would have scored a perfect four out of four. What was your score?

(RELATIVELY COMPLEX) MISCALCULATION

Wason (1968) proposed the following test of formal operational thought. Suppose you are shown four cards with the faces showing respectively "D," "3," "B," and "7," as displayed in the figure below.

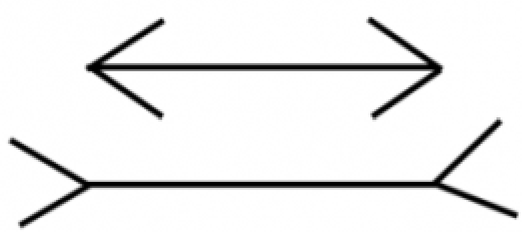
You are told that a card with a number on one side (e.g., 3 or 7) has a letter on the reverse side (e.g., D or B). You are then asked which of the cards you would need to flip over to test the hypothesis that "If there is a D on one side of any card, then there is a 3 on its other side."

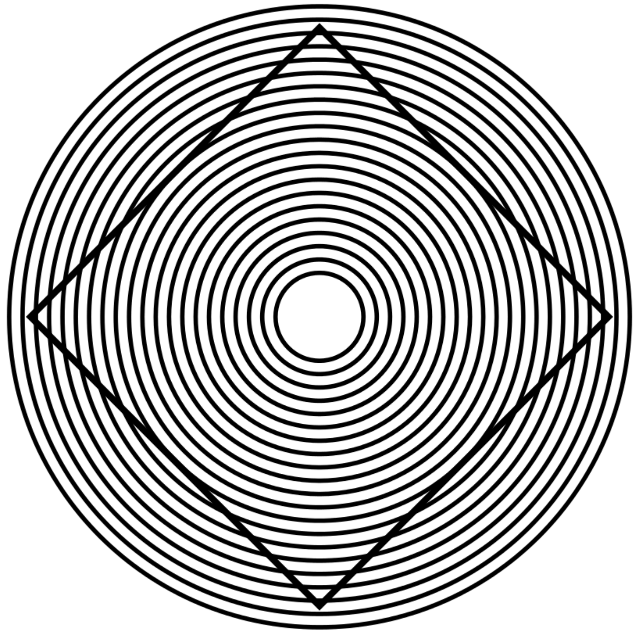
Answer

To test this hypothesis, you would need to flip the D card. However, you would also need to flip over the 7 card as well. If the letter on the opposite side of the 7 card is D, then the hypothesis would be false.

COGNITIVE ILLUSIONS

Ready to be weirded out? Which of the two horizontal lines is the longest?

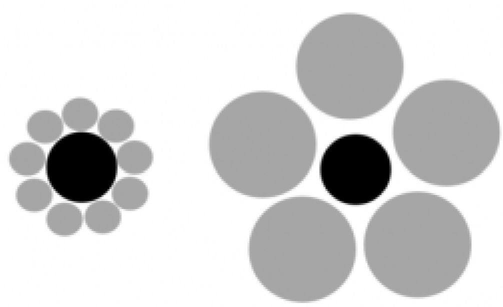




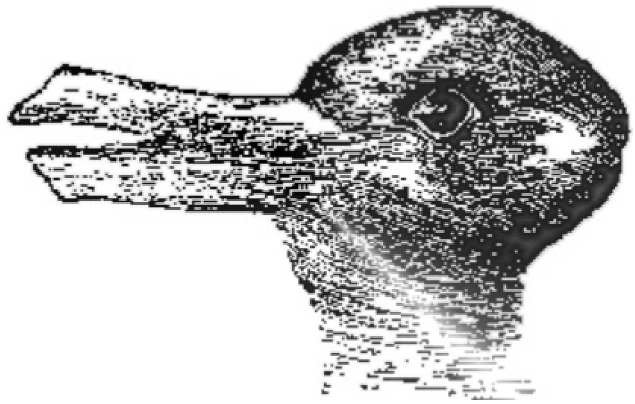
"Ehrenstein", by David Eccles, in the Public Domain

Are the sides of this cube bent inward?

Which of the black circles is the largest?



Do you see a rabbit or a duck in this drawing?



"Rabbit-Duck Illusion", by Fastfission~commonswiki, in the Public Domain

Answers

1. Take a close look. The two lines are identical in length.
2. The sides of the cube are not bent inward.
3. Neither. The black circles are identical in size.
4. You can see both a rabbit and a duck in the drawing.

Homo economicus would provide correct answers to each question, no problem. How did you do? Biederman (1972) and Palmer (1975) contend that our visual perceptions are affected by both our prior conceptual structures and the characteristics of the visual stimulus itself. This would explain why you may have struggled to answer some of questions correctly.

HEURISTICS

A heuristic is a practical, problem-solving method that is not guaranteed to lead to an optimal or rational solution but is nonetheless deemed sufficient by an individual or organization for obtaining a short-term goal or approximation (Myers, 2010). Heuristics can lead *Homo sapiens* to misjudge situations that more reasoned thought or research would otherwise improve upon.[1]

AFFECT HEURISTIC

Have you ever based a decision upon your like or dislike of the object in question rather than on more objective information and logical reasoning? For example, maybe you've based your decision of whether to purchase stock in a company based upon your like or dislike of the company rather than whether the company's stock price is under- or over-valued? If you have ever made a decision like this, then as Kahneman (2011) instructs us, you are guilty of an Affect Heuristic. The key to distinguishing this heuristic is the absence of any information or evidence that might otherwise be used to render judgment or make a decision. We might, therefore, call this the ignorance-is-bliss heuristic.

1. As we will learn in Section 4, heuristics can, in some cases, lead to preferable outcomes.

AVAILABILITY HEURISTIC

Have you ever judged the frequency of an occurrence by the ease with which instances of the occurrence have come to your mind or you have personally experienced it? For example, a judicial error that affected you personally has undermined your faith in the justice system more than a similar incident that you read about in the newspaper? If you have ever judged an occurrence like this, then you were guilty of using an Availability Heuristic.

In an interesting study of the Availability Heuristic, Lichtenstein et al. (1978) asked subjects participating in an experiment whether they knew the likely causes of death in the US. The subjects were told that, on average, 50,000 people die each year due to motor vehicle accidents. They were then asked to state how many people they thought died from 40 other possible causes, ranging from venomous bites or stings, to tornados and lightning strikes, to floods, to electrocution, to fire and flames . . . I think you get the grim picture. The authors found that subjects tended to overestimate the number of people who die from less likely causes and underestimate the number of deaths from more likely causes. For example, the average number of deaths due to fireworks (a less likely cause) was estimated by the experiment's subjects to be over 330 per year when the actual number is only six. And, the number of deaths due to electrocution (a more likely cause) was estimated by the subjects to be roughly 590 versus the actual number of over 1,000.

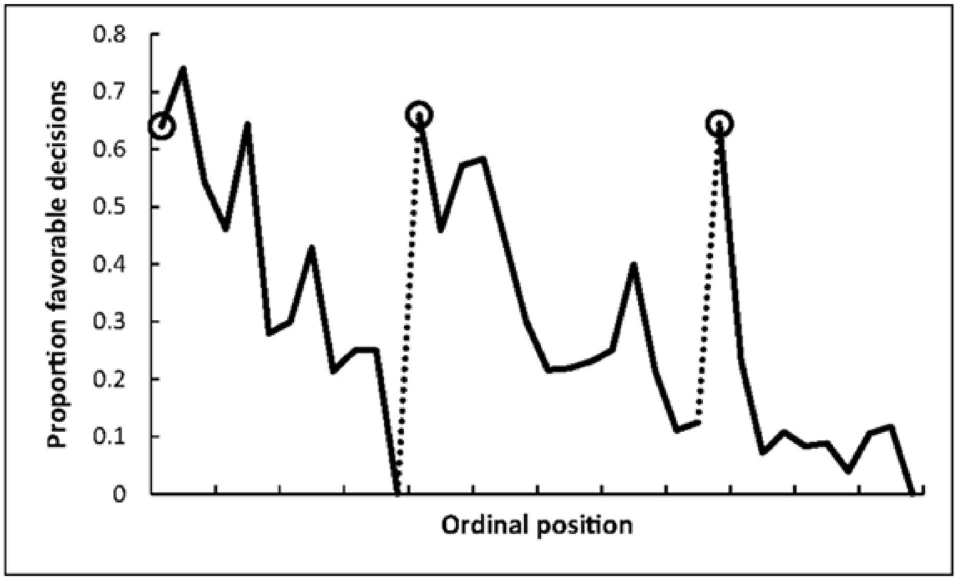
Subjects also tended to believe that two different causes associated with a similar number of deaths were instead associated with markedly different numbers of deaths. For example, homicides and accidental falls account for roughly 18,900 and 17,450 deaths per year, respectively, while, on average, the subjects believed these two death tallies to be roughly 8,440 and 2,600. While the actual ratio of deaths by homicide to deaths by accidental falls is only 1.08 (18,900 ÷ 17,450), the corresponding believed ratio is 3.25 (8,440 ÷ 2,600). Upon further questioning of the subjects, Lichtenstein et al. discovered that this upward bias correlated with newspaper coverage and whether a subject had direct experience with someone who had died from a given cause—the very things that influence an Availability Heuristic.

Needless to say, *Homo economicus* would never deign to use such heuristics. She would be fully informed of the actual death statistics.

EFFECTS

DEPLETION EFFECT

Danziger et al. (2011) studied the proportion of rulings made by parole judges in favor of prisoners' requests for parole. Their results are depicted in the figure below.



(Danziger, et al. 2011)

Circled points in the figure indicate the proportions of first decisions made in favor of parole in each of three decision sessions. The first decision session began after morning break time. The second session began after lunch break, and the third session began after afternoon break time. Tick marks on the horizontal axis denote every third case heard by the judges, respectively, and the dotted lines indicate food breaks.

Note that for each decision session, the rulings begin in favor of parole and then steadily decline as the end of each session is approached. Apparently, the judges get crankier as each session wears on. Their sympathies suffer what's known as a Depletion Effect.

We would expect no such pattern from *Ludex economicus* (judges from the *Homo economicus* species). But what exactly would that pattern be?

PRIMING EFFECT

Consider these two thought experiments.

Last night Sally and Bob went out to dinner together. They enjoyed a meal at Wai Wai's Noodle Palace

S O _ P

Last night Thida came home from work feeling tired and sweaty from a long day of work. She took a long shower.

S O _ P

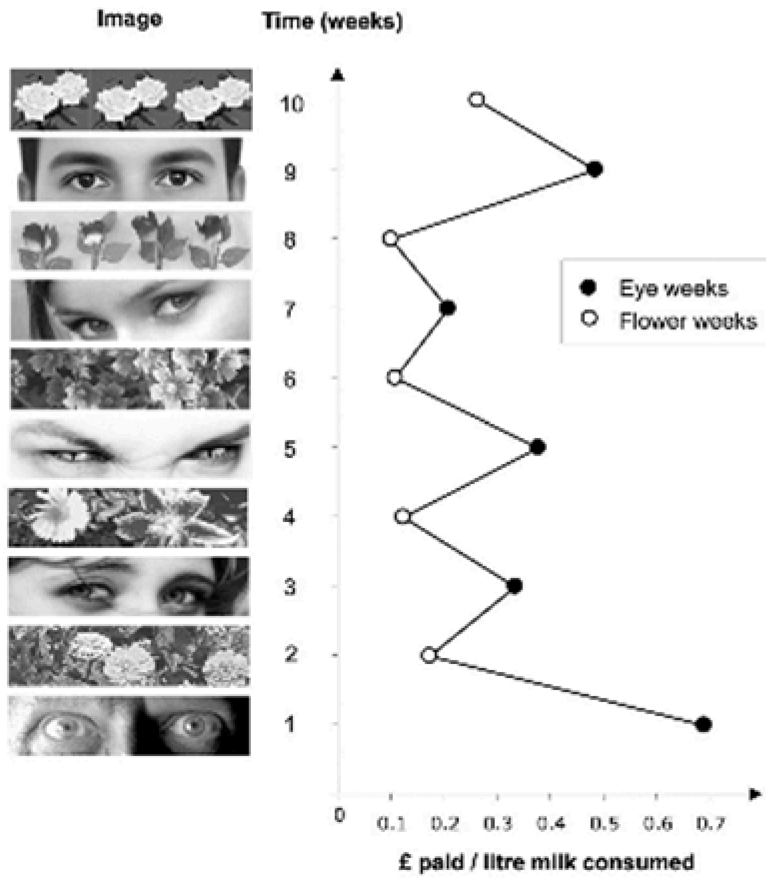
If you chose the letter "u" for the first box and "a" for the second box, then you are likely guilty of what Kahneman (2011) calls a Priming Effect. As these experiments demonstrate, priming *Homo sapiens* is rather easy.

PRIMING EFFECT (VERSION 2)

Bateson et al. (2006) examined the effect of an image of a pair of eyes on contributions made by colleagues to an "honesty box" used to collect money for drinks in a university coffee room. Suggested prices for the drinks were listed as follows ("p" stands for British "pence").

> Coffee (with or without milk): **50p**
>
> Tea (with or without milk): **30p**
>
> Milk only (in your coffee or tea): **10p**
>
> Full cup of milk: **30p**
>
> Please put your money in the blue tin.
>
> Thanks, Melissa.

The figure below presents the study's results.



(Bateson, et al. 2006)

Relative to the week before, when a photo of a floral arrangement was shown, the week associated with human eyes staring back at Melissa's colleagues resulted in more money contributed to the honesty box. All told, the authors report that their colleagues contributed nearly three times as much for their drinks when a pair of eyes were displayed rather than the floral arrangement. This result suggests the importance of the social cue of being watched (and thus, reputational concerns) on cooperative behavior among humans. It is another example of a Priming Effect.

Homo economicus would not have been swayed by such cues and reputational concerns. Instead, he would have exhibited what's known as free-riding behavior, never or only rarely contributing to the honesty box, irrespective of whether a pair of eyes were glaring or flowers blooming.

PRIMING EFFECTS ABOUND

Examples of the Priming Effect abound. For example, Kahneman (2011) mentions research suggesting that "money-primed" people demonstrate more individualism (i.e., more independent-minded and selfish behaviors and a stronger preference for being alone). Berger et al. (2008) find that support for ballot propositions to increase funding for public schools is significantly greater when the polling station is located in a school rather than a nearby location.

Can you identify ways in which you are primed in your daily life? Of course, *Homo economicus* would be compelled to answer "no" to this question.

MERE EXPOSURE EFFECT

Zajonc and Rajecki (1969) ran an interesting field experiment on the campuses of the University of Michigan and Michigan State University. For a period of 25 days, an ad-like box appeared on the front pages of the student newspapers containing one of the following Turkish words: KADIRGA, SARICIK, BIWONJNI, NANSOMA, IKTITAF. The frequency with which the words were repeated varied. One of the words was shown only once, and the others appeared on two, five, ten, or twenty-five separate occasions. No explanations were offered to the readers of the papers. When the mysterious ads ended, the investigators sent questionnaires to readers asking for their impressions of whether each of the words "means something good or something bad." The words presented more frequently were rated much more favorably than the words shown only once or twice. This has come to be known as the Mere Exposure Effect.

INTENTIONAL CAUSATION

Consider the following thought experiment.

> Read this sentence:
>
> After spending a day exploring beautiful sights in the crowded streets of New York City, Jane discovered that her wallet was missing.

What comes to mind? Any chance that Jane was pickpocketed? If so, then you have succumbed to what Kahneman (2011) calls Intentional Causation.

JUMPING TO CONCLUSIONS

The website Effectiviology provides several examples of how we *Homo sapiens* jump to conclusions (https://effectiviology.com/jumping-to-conclusions/). Have you ever made 'jumps' like these?

- Immediately deciding that a restaurant's food is bad because its windows are smudged.
- Believing someone is rich because she drives a fancy car.
- Believing you will fail a test because you struggled with some of the practice questions.
- Thinking someone does not like you because they were not enthusiastic when you said "good morning".
- Thinking a house is on fire because you see smoke coming out of a window.
- Assuming that because you did not get along with one person from a certain social group, you will not get along with anyone else from that group either.

If so, then join the proverbial club. Jumping to conclusions is an easy thing to do.

FRAMING EFFECT

Consider the following thought experiment.

> Different ways of presenting the same information evoke different interpretations. Consider two car owners who seek to reduce their costs:
>
> Sylvester switches from a gas-guzzler of 5 miles per gallon (mpg) to a slightly less voracious guzzler that runs at 6 mpg. The environmentally virtuous Elizabeth switches from a 13 mpg car to one that runs at 17 mpg. Both Sylvester and Elizabeth drive their cars 16,000 miles per year. Who will save more gas by switching?

If you chose Elizabeth you have fallen victim to what is known as a Framing Effect. Guess again.

Elizabeth saves (16,000÷13) – (16,000÷17) = 1,231 – 941 = 290 gallons per year, while Sylvester saves (16,000÷5) – (16,000÷6) = 3,200 – 2,667 = 533 gallons per year!

HALO EFFECT

Consider the following thought experiment.

> Who do you think has more virtuous qualities, **Abigal** or **Anne**?
>
> **Abigal**: intelligent, industrious, impulsive, critical, stubborn, envious
> **Anne**: envious, stubborn, critical, impulsive, industrious, intelligent

Note that Abigal and Anne share the same qualities. The only difference is that the more virtuous qualities are listed first for Abigal and last for Anne. As a result, you are more likely to choose Abigal as having the more virtuous qualities simply because of a type of Framing Effect called the Halo Effect. Surely, *Homo economicus* would have identified Abigal and Anne as equally virtuous individuals.

In one of the earliest laboratory experiments designed to measure the Halo Effect, Nisbett and Wilson (1977) had a group of students observe videotaped interviews with a professor who spoke with a pronounced foreign accent and then rate his "likeability."[2] As the authors point out, when we like a person, we often assume that those attributes of the person about which we know relatively little are also favorable. For example, a person's appearance may be perceived as more attractive if we like the person than if we do not.

The subjects in Nisbett and Wilson's experiment (roughly 120 University of Michigan students enrolled in an introductory psychology course) were told that the investigators were studying the

2. This was by no means the first such Halo-Effect experiment. For instance, an earlier experiment conducted by Landy and Sigall (1974) found that evaluations of an essay (written by an unknown author) made by male college students were graded substantially higher when the alleged author was an attractive woman rather than an unattractive woman. This Halo Effect was pronounced, especially when the essay was of relatively poor quality.

possibility that ratings of an instructor presented in such a brief fashion might resemble ratings by students who had taken an entire course with the instructor. The subjects were shown one of two different seven-minute, videotaped interviews with the same instructor, a native French-speaking Belgian who spoke English with a fairly pronounced accent. In one interview, the instructor presented himself as a likable person, respectful of his students' intelligence and motives, flexible in his approach to teaching, and enthusiastic about his subject matter (i.e., he portrayed himself as a "warm teacher"). In the other interview, the instructor appeared to be quite unlikable, cold and distrustful toward his students, rigid and doctrinaire in his teaching style (i.e., portraying a "cold teacher"). After viewing the videotaped interview, the subjects rated the instructor's likability, as well as the attractiveness of his physical appearance, his mannerisms, and his accent. It was anticipated that the subjects would rate the instructor as having a more attractive physical appearance, more attractive mannerisms, and a more attractive accent when he was likable than when he was unlikable.

A substantial majority of the subjects who observed the interview with the warm teacher rated his physical appearance as appealing, whereas a substantial majority of those who observed the interview with the cold teacher rated his appearance as irritating. Similarly, a majority of subjects viewing the warm teacher rated his mannerisms as appealing, whereas a majority of subjects viewing the cold teacher rated his mannerisms as irritating. Lastly, about half of the subjects viewing the warm teacher rated his accent as appealing, while half rated the accent as irritating, whereas the overwhelming majority of subjects who viewed the cold teacher rated his accent as irritating.[3]

Hence, it appears that unlike *Homo economicus*, who would not be swayed by inconclusive evidence such as a seven-minute interview, *Homo sapiens* can indeed be influenced by these types of first encounters and attendant impressions. We tend to fall prey to the Halo Effect.

ORDERING EFFECT

Hogarth and Einhorn (1992) investigate what's known as the Ordering Effect, which is associated with how *Homo sapiens* update their beliefs over time—for example, how first impressions of an acquaintance are updated as you spend more time together.[4] The authors consider three pertinent questions concerning this updating process. First, under what conditions does information processed earliest in the updating sequence have greater influence (i.e., produce a Primacy Effect)? Second, under what conditions is later information more important (i.e., produce a Recency Effect)? And third, under what conditions is order irrelevant? In general, Hogarth and Einhorn consider order effects of the following type:

> There are two pieces of evidence, A and B. Some subjects express an opinion after seeing the information in the order A-B; others receive the information in the order B-A. An order effect occurs when opinions formed after A-B differ from those formed after B-A.

To test for Primacy and Recency Effects, the authors present subjects in their experiments with a set of four scenarios, each of which involves an initial description (the stem) and two additional pieces of information presented in separate paragraphs (the evidence). The content of the four stems consists of the following: (1) a defective stereo speaker thought to have a bad connection; (2) a baseball player named Sandy whose hitting has improved dramatically after a new coaching program; (3) an increase in sales of a supermarket product following an advertising campaign; and (4) the contracting of lung

3. In Chapter 5 we will learn how researchers discern differences like these on a more formal, statistical basis.

4. Similar to how the Halo Effect represents a special case of a Framing Effect, you should recognize that the Ordering Effect is likewise a special case of a Framing Effect.

cancer by a worker in a chemical factory. Note that each stem consists of an outcome (e.g., Sandy's hitting has improved dramatically), and a suspected causal factor (e.g., a new coaching program). After reading a stem, subjects are asked to rate how likely the suspected causal factor was the cause of the outcome on a rating scale from 0 to 100. For example, in the baseball scenario, subjects are asked, "How likely do you think that the new training program caused the improvement in Sandy's performance?"

In one experiment (Experiment 1), subjects are provided with both "strong" and "weak" positive evidence to nudge them toward a revised answer. Continuing with the baseball scenario, the positive evidence consists of two sentences: "The other players on Sandy's team did not show an unusual increase in their batting average over the last five weeks. In fact, the team's overall batting average for these five weeks was about the same as the average for the season thus far." The first sentence provides strong positive evidence and the second sentence provides weak positive evidence. Thus, the evidence is provided in a "strong-weak order" (strong-weak and weak-strong orderings were randomized across subjects). After reading the evidence, subjects are asked again to rate how likely the suspected causal factor was the cause of the outcome on a rating scale from 0 to 100.

In Experiments 2 and 3, the same procedures were followed except that in Experiment 2 the two pieces of evidence consist of strong negative and weak negative information about the outcome, and in Experiment 3 the two pieces of information are mixed, involving positive and negative information. An example of negative information in the baseball scenario is, "The games in which Sandy showed his improvement were played against the last place team in the league. Pitchers on that team are very weak and usually allow many hits and runs."

Hogarth and Einhorn's hypotheses were that subjects participating in Experiments 1 and 2 should not exhibit an Ordering Effect since the evidence was either purely positive or purely negative—the ordering of strong vs. weak should, therefore, not measurably impact a subject's initial rating of the likelihood of the suspected causal factor having caused the outcome. However, the ordering of the mixed evidence in Experiment 3—positive-negative vs. negative-positive—should impact a subject's initial rating.

Each of the authors' hypotheses was confirmed by the experiments. In Experiment 3 they found statistically significant evidence of a Recency Effect.[5] Specifically, the positive-negative order resulted in an average decrease in the subjects' ratings of slightly more than 9, relative to the average initial judgment, and the negative-positive order resulted in an average rating increase of slightly less than 3. Recency in this case is tied to the evidence provided in the second sentence as opposed to the first sentence. Had the result been reversed (i.e., it was the first sentence that drove the average change in rating rather than the second sentence), then Hogarth and Einhorn would have instead found evidence of a Primacy Effect.

Of course, we would expect neither recency nor primacy to affect *Homo economicus*.

ANCHORING EFFECT

Kahneman (2011) describes another effect known as the Anchoring Effect, whereby a subject's answer to a question is anchored to information that is contained in the question itself. For example, suppose Individual 1 is presented with Question 1 below, and Individual 2 is presented with Question 2.

5. We explicitly define what we mean by "statistically significant" in Section 4. For now, think of statistically significant this way: the result of an experiment is statistically significant if it is likely not caused by chance for some given level of confidence, typically 95%.

Assume that both individuals are so alike we can almost think of them as clones of one another. Neither of them actually knows how old Gandhi was at death.

1. Was Gandhi younger or older than 114 years at his death? How old was Gandhi at his death?
2. Was Gandhi younger or older than 35 years at his death? How old was Gandhi at his death?

If the two individuals each suffer from the Anchoring Effect, then Individual 1 will answer a higher age than Individual 2. This is because Individual 1's anchor age in his or her question, 114 years, is so much higher than Individual 2's anchor of 35 years. Based on their disparate answers, an Anchoring Index can be calculated as (Individual 1's answer – Individual 2's answer) ÷ (114 – 35).

Of course, if the two individuals happen to be from the species *Homo economicus*, they would both answer 78 years old, which was Gandhi's actual age at death. And in this case, their Anchoring Index would equal zero!

In a classic test of the anchoring effect among *Homo sapiens*, Ariely et al. (2003) asked students in a laboratory experiment whether they would be willing to purchase a box of Belgian Chocolates for more money than the last two digits of their Social Security Numbers (SSNs). For example, if the last two digits of a participant's SSN were 25, then s/he was asked whether s/he would be willing to pay (WTP) more than $25 for the chocolates. The participants were then asked for the specific amount they would be WTP. Because SSNs are assigned randomly, the authors hypothesized that there should be no relationship between the participants' SSNs and their respective WTP values. On the contrary, Ariely et al. (2003) found a positive relationship between the participants' SSNs and WTP values, suggesting that a *Homo sapiens*' SSN can induce an Anchoring Effect, particularly when it comes to our WTP values for Belgian Chocolates. Yum!

In a separate experiment, Ariely et al. sought to answer the attendant question, do *Homo sapiens* flip-flop from one anchor price to another, continually changing our WTP values? Or does the first anchor price we encounter serve as our anchor over time and across multiple decisions (i.e., do we exhibit what the authors call coherent arbitrariness)? For their experiment, the authors recruited approximately 130 students attending a job recruitment fair on the MIT campus. The experiment subjected each participant to three different sounds through a pair of headphones. Following each sound, the participants were asked if they would be willing to accept a particular amount of money (which served as the experiment's anchor price) for having to listen to the sounds again. One sound was a 30-second, high-pitched, 3,000-hertz sound, mimicking someone screaming in a high-pitched voice (Sound 1). Another was a 30-second, full-spectrum (white) noise, similar to the noise a television set or radio makes when there is no reception (Sound 2). The third was a 30-second oscillation between high-pitched and low-pitched sounds (Sound 3). Ariely et al. used these particular sounds due to there being no existing market for annoying sounds (therefore, the participants were precluded from confounding their responses in the experiment with a pre-existing market price).

For the first part of the experiment, anchor prices of 10 cents or 90 cents were randomly assigned to the participants. After indicating whether they would accept their anchor price for listening to Sound 1 again ("yes" or "no"), each participant then indicated the lowest price they would willingly accept to listen to the sound again. Participants whose price was lowest "won" the opportunity to hear Sound 1 again, and actually got paid for doing so. The remaining participants were not given the opportunity to listen to the sound again (and thus, were not paid for this part of the experiment). As expected, the authors found that those participants whose anchor price was 10 cents stated a lower willingness to listen value to Sound 1 again (33 cents on average) relative to those whose anchor price had been 90 cents (73 cents on average).

To test how influential the anchor prices of 10 cents and 90 cents were in determining future decisions, Ariely et al. then subjected each participant (from both the 10-cent and 90-cent anchor price groups) to Sound 2 and asked if they would be willing to accept a payment of 50 cents to endure the sound again. Similar to the first part of the experiment, after indicating whether they would accept 50 cents for listening to Sound 2 again, each participant stated the lowest price they would willingly accept to listen to the sound again. It turned out that the original 10-cent group stated much lower prices than the original 90-cent group. Although both groups had subsequently been exposed to the 50-cent anchor price, their original anchor prices (10 cents for some, 90 cents for others) predominated. In other words, *Homo sapiens* exhibit persistent Anchoring Effects.

In the experiment's final stage, participants were instructed to listen to Sound 3. This time, Ariely et al. asked each of the original 10-cent group members if they would be willing to listen to this sound again for 90 cents. And Ariely et al. asked each of the original 90-cent group members if they would be willing to listen to this sound again for 10 cents. Having flipped the anchor prices, the authors could now discern which anchor price—the first or the second—exerted the greatest influence on the participants' stated prices. Once again, each participant was then asked how much money it would take to willingly listen to Sound 3 again.

The final results were that (1) those participants who had first encountered the 10-cent anchor price stated relatively low prices to endure Sound 3 again, even after 90 cents was stated as the subsequent anchor price, and (2) those who had first encountered the 90-cent anchor price demanded relatively high prices, even after 10 cents was stated as the subsequent anchor price. Therefore, Ariely et al. conclude that our initial decisions anchor future decisions over time. Or, to put it another way, first impressions are important. Anchoring Effects remain with us long after an initial decision is made. This is what explains, for example, the heuristic of brand loyalty. As Ariely (2008) points out, loyal Starbucks customers likely share the same story explaining their fealty. Following their first experience drinking a Starbucks coffee, they apply the following heuristic: "I went to Starbucks before, and I enjoyed both the coffee and the overall experience, so this must be a good decision for me." And so on. This can also explain how you might start with a small drip coffee (your anchor) and subsequently work your way up to a large Frappuccino.[6]

SILO EFFECT

A Silo Effect occurs when a system is not in place that enables separate departments or teams within an organization to communicate effectively with each other. Productivity and collaboration suffer as a result. A classic example of the Silo Effect is when two departments within a given organization are working on practically identical initiatives or projects but neither department is aware of what the other is doing (Marchese, 2016). This phenomenon is also known as homophily (i.e., when contact occurs more often between similar than dissimilar departments).

Although conventional wisdom has suggested for some time now that breaking down silos and fostering interorganizational partnerships to achieve public health outcomes has distinct advantages,

6. Ariely goes on to explain what likely attracted you to Starbucks in the first place.

and examples can indeed be found where the practice of collaboration is growing within the public health system, Bevc et al. (2015) set out to measure the extent to which disciplinary and organizational silos that have traditionally characterized public health still exist. In particular, the authors test for the persistence of Silo Effects in over 160 public health collaboratives (PHCs); social networks comprised of diverse types of partners (e.g., including law enforcement agencies, nonprofit advocacy groups, hospitals, etc.); varying levels of interaction; and multiple configurations designed to increase common knowledge and resource sharing. Interestingly, Bevc et al. find that as network size increases, a potential bias is observed among specific organization types in terms of their choosing to interact with similar organizations (e.g., for law enforcement agencies to collaborate with other law enforcement agencies, nonprofits with other nonprofits, and public health organizations with other public health organizations, etc).

Thus, even in settings where reducing the impulse for homophily is explicitly being targeted, *Homo sapiens* persist in occupying their silos. Given their ubiquitous understanding of the benefits of collaboration among dissimilar groups, *Homo economicus* would never have built such silos in the first place.

STUDY QUESTIONS

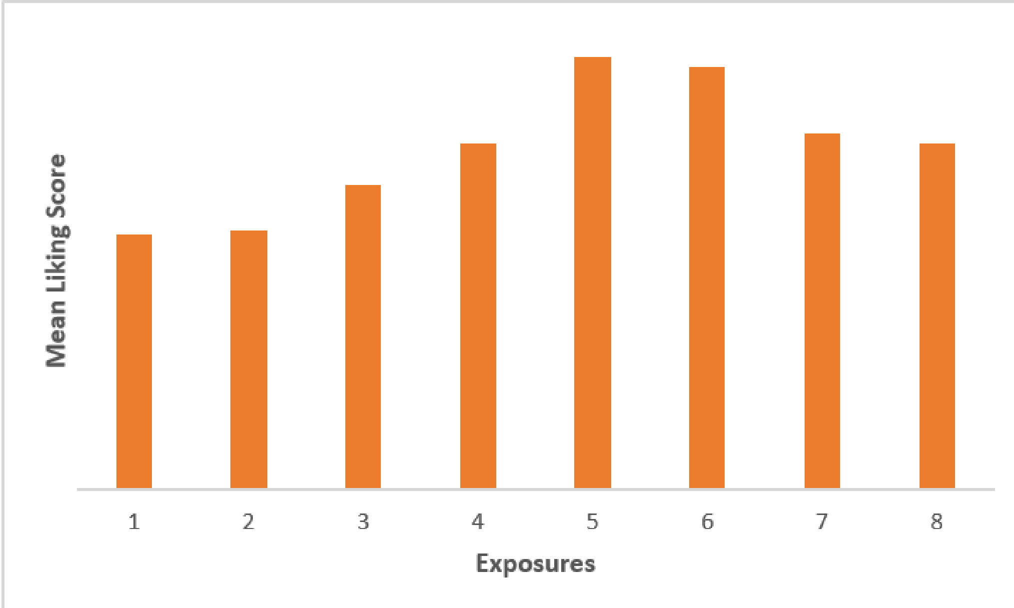
1. Describe two instances in your own life where you have adopted the Affect and Availability Heuristics to help you in making decisions. What drove you to adopt these heuristics? Do you believe the heuristics served you well? Why or why not?

2. Browse the internet for a challenge question that, like those presented in this chapter, instigate miscalculation and error in reasoning. Also, find a cognitive illusion that elicits the same sense of wonderment as those presented in this chapter.

3. Can you think of another sector of society besides the judiciary where the Depletion Effect has potentially profound implications? Explain.

4. The Honesty Box described in Priming Effect (Version 2) is an example of a public good funded by voluntary contributions, and the human eye-floral arrangement prompts are pre-contribution mechanisms designed to induce full payment by coffee-room attendees. Can you think of a post-contribution mechanism that might also induce full payment to the Honesty Box? How would this mechanism actually work?

5. Suppose you have conducted a field experiment with a group of 50 adults to measure the incidence of a Mere Exposure Effect. You have them listen to the new Bruce Springsteen song “Letter to You” once per day over a period of eight consecutive days, and then register their Liking Score (the extent to which they have enjoyed listening to the song) after each listen. You summarize your results in the bar graph below. Are these results evidence of a Mere Exposure Effect?



6. **Warning**: This question concerns a politically charged event that occurred on January 18, 2019, at the Indigenous People's March in Washington, D.C. After reading this account of what happened at the march, and viewing this video of the event, which of the effects presented in this chapter do you think best describes this episode in our nation's history?

7. Think of a situation in your own life when you framed information (either wittingly or unwittingly) in such a way that helped pre-determine an outcome. Describe the situation and how you framed the information. Was the outcome improved or worsened as a result of how you framed the information?

8. After having learned about the Anchoring Effect in this chapter, do you think you will ever fall for something like this again?

9. When someone admonishes you "not to judge a book by its cover," or as British management journalist Robert Heller once noted, "Never ignore a gut feeling, but never believe that it's enough," what heuristic(s) is he unwittingly advising you to avoid using?

10. Browse the internet for information about an effect that was not discussed in this chapter. Can you classify this effect as a special case of a Priming or Framing Effect? Explain.

11. Browse the internet for a heuristic other than the Affect and Availability Heuristics described in this chapter. Explain the heuristic.

12. It's one thing to detect the existence of a Silo Effect and quite another to measure its

negative impacts on relationships between organizations or individuals. Identify a setting or situation where a Silo Effect exists and design a field experiment to measure the impacts of this effect on an outcome of interest.

13. The Halo Effect suggests that someone who is perceived as being physically attractive has an advantage in certain situations—for example, when applying for a job. Can you think of why the halo might have a reverse effect?

Media Attributions

- Figure 6 (Chapter 1) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 8 (Chapter 1) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 8a (Chapter 1) © David Eccles is licensed under a Public Domain license
- Figure 8b (Chapter 1)
- Rabbit-Duck Illusion © Fastfission~commonswiki is licensed under a Public Domain license
- Figure 12 (Chapter 1) © 2021 National Academy of Sciences is licensed under a All Rights Reserved license
- Figure 12 (Chapter 1) © 2006 The Royal Society is licensed under a All Rights Reserved license
- Figure 17 (Chapter 1) © Arthur Caplan is licensed under a CC BY (Attribution) license

CHAPTER 2.

THE BIASES AND FALLACIES OF *HOMO SAPIENS*

The quirks discussed in Chapter 1 set the stage for biases and fallacies that often plague the choice behavior of *Homo sapiens*. The extent to which we miscalculate problems—problems begging for various degrees of logical reasoning—illuminates our inherent cognitive limitations. The effects that prime and frame us to miscalculate and misjudge underlying conditions (or when harnessed for our betterment, help us correct our otherwise misguided thought processes) evince the different contexts within which our quirks lead us astray. As we learned in Chapter 1, *Homo sapiens* have devised heuristics, or rules-of-thumb, that we use as substitutes for deeper analysis of, or reasoning about, problems that otherwise warrant such depth and reasoning.

In this chapter, we make a subtle turn from these stage-setting quirks toward identifying the different ways in which *Homo sapiens* are innately biased. The word "innately" is important here. Bear in mind that the biases discussed in this chapter are, for the most part, ingrained in the human condition. They are not learned prejudices.

STATUS QUO BIAS

As previously discussed in Example 3 in the Introduction, *Homo sapiens* are prone to what's known as Status Quo Bias, whereby we tend to resist change, and thus, sometimes miss opportunities to make beneficial changes in our personal lives as well as those for society at large. In Example 3 we learned that choosing the default option carefully can solve the problem of status quo bias. For instance, the nations of Spain, Belgium, Austria, and France have among the highest organ donation consent rates. Why? They use "opt-out" default on the organ donation registration form, and thus, presume consent on the behalves of their citizens (Davidai et al., 2012). Similarly, Thaler and Benartzi (2004) proposed the Save More Tomorrow (SMarT) Program where employees commit in advance to allocate a portion of their future salary increases toward retirement savings (in effect opting out of the alternative of not making this allocation). The authors found that (1) 78% of the company's employees offered the plan joined, (2) 80% remained in the program for 40 months, and (3) the company's average savings rate increased by 10%.

CONFIRMATION BIAS

In our daily lives, we sometimes guard more against committing what Statisticians call Type 2 error (failing to reject a false null hypothesis) than against committing Type 1 error (rejecting a true null hypothesis). For example, one of my most frequent Type 2 errors is instinctively blaming my wife whenever something of mine goes missing at home, like a pair of shoes or a magazine. In this case, I effectively set the null hypothesis as "my wife is innocent." In my mind though, I consider her guilty. I am therefore afraid of failing to reject the null hypothesis that she is innocent. Further, I unwittingly guard my ego against being proved wrong by sharing menacing looks and grimaces and sighs of helplessness. This is my guard against committing what I suspect would be a Type 2 error.

Unfortunately, this precautionary tendency leads me to commit what is known as Confirmation Bias.[1] The way I structure my thinking, and the way I behave, impel me to confirm to myself that I am not absent-minded. It's my wife's fault. She is the guilty one!

Darley and Gross (1983) conducted one of the earliest and most enduring studies of Confirmation Bias with approximately 70 of their undergraduate students. One subgroup of the students was subtly led to believe that a child they were observing came from a high socio-economic (SE) background, while another subgroup was subtly led to believe that the child came from a low SE background. Nothing in the child's SE demographics conveyed information directly relevant to the child's academic abilities. When initially asked by the researchers, both subgroups rated the child's ability to be approximately at grade level. Two other subgroups, respectively, received specific SE demographics about the child—one set of demographics indicating that the child came from a high SE background, the other that she came from a low SE background. Each of these two subgroups then watched the same video of the child taking an academic test. Although the video was identical for all students in these two subgroups, those in the subgroup who had been informed that the child came from a high SE background rated her abilities well above grade level, while those in the subgroup for whom the child was identified as coming from a lower SE background rated her abilities as below grade level. The authors concluded that *Homo sapiens* are prone to using some "stereotype" information to form hypotheses about the stereotyped individual. *Homo sapiens* test these hypotheses in a biased fashion, leading to their false confirmation.

LAW OF SMALL NUMBERS

Consider the following thought experiment presented in Kahneman (2011):

> Consider three possible sequences for the next six babies born at your local hospital (**B** stands for "boy" and **G** stands for "girl"):
>
> **B B B G G G**
>
> **G G G G G G**
>
> **B G B B G B**
>
> Are these sequences equally likely?

Kahneman's guess is that you answered "no," in which case you consider the third sequence (**B G B B G B**) to be the most likely to occur. You liken it to your experience flipping a coin—flip the coin often enough and you would expect to see Heads and Tails alternating more repeatedly, which would be a valid expectation. The problem here is that the coin hasn't yet been flipped often enough. A sequence of six babies is not enough flips of the coin, so to speak, to necessarily start witnessing an alternating

1. Of course, it can be argued that those who guard more against committing Type 1 error are just as likely to exhibit Confirmation Bias. In my case, this counterfactual situation would have me guarding against blaming my wife. But since our children no longer live under our roof, and we own no pets capable of playing hide and seek with my things, this scenario would require that I blame myself first and foremost. Perish the thought!

pattern. Hence, if you indeed answered "no" to the question you are guilty of what's known as the Law of Small Numbers.[2] Needless to say, *Homo economicus* is not beholden to this law.

In a classic study of the Law of Small Numbers, Gilovich, et al. (1985) investigated common beliefs about "the hot hand" and "streak shooting" in the game of basketball. The authors found that basketball players and fans alike tend to believe that a player's chance of hitting a shot is systematically greater following a hit than following a miss on the previous shot. However, data compiled for an entire season (i.e., data from a large sample) did not support the hot-hand hypothesis which, by its very nature, is predicated on a game-by-game basis (i.e., data from a small sample).

In specific, Gilovich, et al. (1985) found that if a given player on a given night had just missed one shot, then on average, he hit 54% of his subsequent shots. Likewise, if the player had just hit one shot, he hit 51% of his remaining shots. After hitting two shots, he then hit 50% of subsequent shots. The estimated correlation coefficient between the outcome of one shot and the next was a statistically insignificant –0.039, suggesting that shooting streaks are an illusion. Each shot is essentially independent of the previous shot. To the contrary, when surveyed, basketball fans on average expected a 50% shooter to have a 61% chance of making a second shot once the first was made. The authors also analyzed game-by-game shooting percentages to see if a player's performance in a single game could be distinguished from any other game. Again, contrary to the beliefs of the average fan, they found no evidence that players have hot and cold shooting nights.[3]

REPRESENTATIVE BIAS

Consider the following two-part thought experiment presented in Kahneman (2011):

> James is a new student at your university. What is the likelihood that James' major is?
>
> **Psychology Philosophy Chemistry Computer Sciences Library and Info Sciences Physics**

Hopefully, you appealed to some base-rate information in answering this question, either because you recently happened to see some published statistics about your university's distribution of majors

2. The opposite of this law, the Law of Large Numbers, is what underpins the correct answer, "yes," to this experiment's question. Interestingly, the law's application here does not specify a threshold number of births beyond which we would expect to see more of an alternating pattern. Rather, as applied here, the Law of Large Numbers merely implies that six is too small a number.

3. Clotfelter and Cook (1991) and Terrell (1994) tested for a version of the Law of Small Numbers known as the Gambler's Fallacy based upon state lottery data from Maryland and New Jersey. In both lotteries, players try to correctly guess a randomly drawn, three-digit winning number. Both studies found that relatively few players bet on a number that had recently won the lottery. Gamblers in the New Jersey lottery who succumbed to this gambler's fallacy paid more of a price than those who succumbed in the Maryland lottery. This is because in the Maryland lottery all players who pick the correct number win the same prize amount, while in New Jersey a jackpot amount is split evenly among all the winners. Thus, a player in the New Jersey lottery wins more the fewer the number of other winners, in which case picking a number that recently won the New Jersey lottery is actually a better strategy (or less-worse strategy) in New Jersey than it is in Maryland. One would therefore expect to find fewer players succumbing to the Gambler's Fallacy in New Jersey than in Maryland. Although this did occur, the difference was only slight, suggesting that gamblers are hard-pressed to overcome the fallacy even when it is in their best interest to do so.

or because you've wondered about this very question before and, based upon your perceptions, can guesstimate a fairly reasonable answer.

Now, consider a slightly altered form of this experiment.

> James is a new student at your university. During his senior year of school his school's psychologist made the following personality sketch of James based upon tests of uncertain validity:
>
> "James is of high intelligence. He has a need for order and clarity, for neat and tidy systems in which every detail finds its appropriate place. His writing is rather dull and mechanical. He has a strong drive for competence. He seems to have little sympathy for other people and does not enjoy interacting with others. Self-centered, James nonetheless has a deep moral sense."
>
> What is the likelihood that James' major is?
>
> **Psychology** **Philosophy** **Chemistry** **Computer Sciences** **Library and Info Sciences** **Physics**

Kahneman's guess is that the added information provided by James' high school psychologist has made it more likely that your ranking of the likelihoods of each candidate major looks something like this.

Chemistry Physics Computer Sciences Philosophy Psychology Library and Info. Studies

In which case you would be guilty (justifiably so?) of what's known as Representative Bias.

CONJUNCTION FALLACY

Consider the following thought experiment proposed by Kahneman (2011):

> Ella is 31 years old, single, outspoken, and very bright. She majored in Philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in the 2017 Women's March in Washington, DC.
>
> Referring to the list below (which is presented in no particular order), rank each statement from the most to least likely:
>
> 1. Ella is a teacher in a primary school.
> 2. Ella works in a bookstore and practices yoga.
> 3. Ella is active in promoting women's rights.
> 4. Ella is a social worker.
> 5. Ella is a member of Equality Now.
> 6. Ella is a bank teller.
> 7. Ella is an insurance agent.
> 8. Ella is a bank teller and is active in promoting women's rights.

If you ranked number 8 higher than numbers 3 or 6, then you are guilty of a Conjunction Fallacy. That is because numbers 3 and 6 are each marginal probabilities and number 8 is a corresponding joint probability. By definition, marginal probabilities are never less than corresponding joint probabilities. But not to worry. Kahneman reports that in repeated experiments, over 80% of undergraduate and graduate students at Stanford University make the mistake. *Homo economicus*? Never.

CONJUNCTION FALLACY (VERSION 2)

As another example of the Conjunction Fallacy, consider the following thought experiment provided by Kahneman (2011):

> Suppose I have a six-sided die with four green faces and two red faces, which will be rolled 20 times. I show you three sequences that could potentially arise during any subset of the 20 rolls (G = green and R = red):
>
> 1. R G R R R
> 2. G R G R R R
> 3. G R R R R R
>
> Choose the sequence you think is most likely to have arisen during a subset of 20 rolls.

If you chose Sequence 2, you have unwittingly fallen victim to a Conjunction Fallacy. Note that Sequence 1 is a subset of Sequence 2. Similar to the marginal vs. joint probability comparison in the previous thought experiment, subsets of a larger set are always more likely to occur than the larger set itself. To derive this result mathematically, begin by noting that the probability of a red face occurring on any given roll of the die is 2/6 = 1/3 and the probability of a green face is 4/6 = 2/3. Since the outcome of each roll of the die is independent from the outcomes of any other roll, the probability of Sequence 1 is therefore 1/3 x 2/3 x 1/3 x 1/3 x 1/3 = 0.00823, and the probability of Sequence 2 is 2/3 x 0.00823 = 0.00549.

Guess who would have done this math in his or her head, and thus, never chosen Sequence 2?

PLANNING FALLACY

Kahneman (2011) recounts an anecdote about a company's management team that had unfortunately developed over time a systematic tendency toward unrealistic optimism about the amount of time required to complete any given project, as well as the project's probable outcome. Unrealistic optimism is a symptom of what is known as Planning Fallacy, which in turn results in Optimism Bias. As Kahneman (2011) informs us, one way to eschew this bias is to conduct a "premortem" before a project begins, whereby the management team imagines that the project has failed and then works backward (a technique known as backward induction, which you will learn more about in Section 3 of this textbook) to determine what could have potentially lead to the project's failure. This strategy seems to align with the old adage, "hope for the best, expect the worst."

STEREOTYPING

Consider the following thought experiment.

A Taxi in Yangon (the capital city of Myanmar) was involved in a hit-and-run accident at night. Two taxi companies, Grab and Hello, operate in the city.

You are given the following data:

- The two companies operate the same number of taxis, but Grab taxis are involved in 85% of accidents.
- A witness identified the taxi as Hello. The court tested the reliability of the witness under circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two taxis 80% of the time and failed to identify them 20% of the time.

Q: What is the probability that the taxi involved in the accident was Hello rather than Grab?

If you are like most people, you see this experiment as suffering from TMI. In your mind, the great majority of the evidence suggests that a Grab taxi was the culprit. Thus, although you might not assign a probability as low as 100% – 85% = 15% to a Hello taxi having been involved in the accident, chances are you will assign something close to 15%, in which case you are guilty of adopting an Availability Heuristic (recall the discussion about this heuristic in Chapter 1) and consequently stereotyping poor old Grab taxi company. Mathematically, we can use the information supplied in the thought experiment and appeal to what's known as Bayes Rule to calculate Hello's actual probability.

Let $Pr(H)$ = 0.15 represent Hello's (H's) probability of getting in an accident on any given night.

Let $Pr(G)$ = 0.85 represent Grab's (G's) probability of getting in an accident on any given night.

Let $Pr(H|W)$ = 0.80 represent the probability that H was involved in the accident given the witness's (W's) testimony.

Let $Pr(G|W)$ = 0.20 represent the probability that G was involved in the accident given W's testimony.

Via Bayes Rule, the probability that a Hello taxi was involved in the accident is calculated as,

$$[Pr(H|W)\,x\,Pr(H)] \div [(Pr(H|W)\,x\,Pr(H)) + (Pr(G|W)\,x\,Pr(G))]$$

=[0.80 x 0.15] ÷ [(0.80 x 0.15)+(0.20 x 0.85)]=41%,

which (surprise, surprise) is exactly the probability that *Homo economicus* would calculate.

Stereotyping does not materialize solely as a lack of application or misapplication of a mathematical rule. It is a much more pervasive behavior among *Homo sapiens*, particularly when it comes to ascribing motives to or judging the practices of other individuals or groups of people. How do we integrate our impressions of another person to form a perception of that individual's reference group as a whole? To what extent does the Availability Heuristic lead to stereotyping in instances such as these?

Rothbart et al. (1978) preface their experiments in pursuit of answers to these questions with the basic understanding that among *Homo sapiens* information obtained about other individuals is organized mnemonically. Attendant judgments made about the other individuals' respective reference groups in turn vary according to the way the information is organized in one's brain. In particular, when we have repeated interactions with individuals of a specific group, we may organize our perceptions of this group around the specific individuals with whom we have interacted or around an integration of the repeated characteristics of the individuals we have encountered.

Further, Rothbart et al. postulate that one of the strongest determinants of mnemonic organization is the demand made on memory during the learning process. When there is low demand on memory, individuals can organize their perceptions of a group around their interactions with its individual members. However, under a high-memory load, individuals are more apt to organize their perceptions around the integration of repeated characteristics encountered within the group as a whole (ignoring the specific individuals encountered). To investigate these issues of memory, organization, and judgment driving *Homo sapiens*' proclivity to stereotype, the authors designed a series of experiments to examine the effects of memory organization on the recall and heuristic judgments of a reference group's characteristics.

In the experiments, over 200 subjects are presented with identical trait information about group members in one of two ways. In the single-exposure condition, each presentation of a trait is paired with a different "stimulus person," where each stimulus person is encountered only once. In the multiple-exposure condition, a given trait (e.g., "lazy") is paired multiple times with different stimulus persons. Eight favorable traits (cooperative, objective, intelligent, generous, creative, resourceful, sincere, reliable) and eight unfavorable traits (clumsy, anxious, impatient, lazy, compulsive, irritable, withdrawn, stubborn) were used in the experiments. Depending upon the particular experimental session, the same number of desirable traits were presented either one-third as often, half as often, or three-times more often than undesirable traits. The actual experiments themselves consisted of either 16 (low-memory load) or 64 (high-memory load) name-trait pairings.

After the presentation of the stimulus information, subjects estimated the proportions of desirable, undesirable, and neutral persons in the group, recalled the adjective traits, and rated the attractiveness of the group as a whole. Two measures of group attractiveness were obtained by asking subjects to rate the desirability of the group as a whole on a 17-point scale from 1 (extremely undesirable) to 17 (extremely desirable) and to rate "how much they would like a group with these characteristics to be among (their) close friends" on a 17-point scale from 1 (dislike very much) to 17 (like very much).

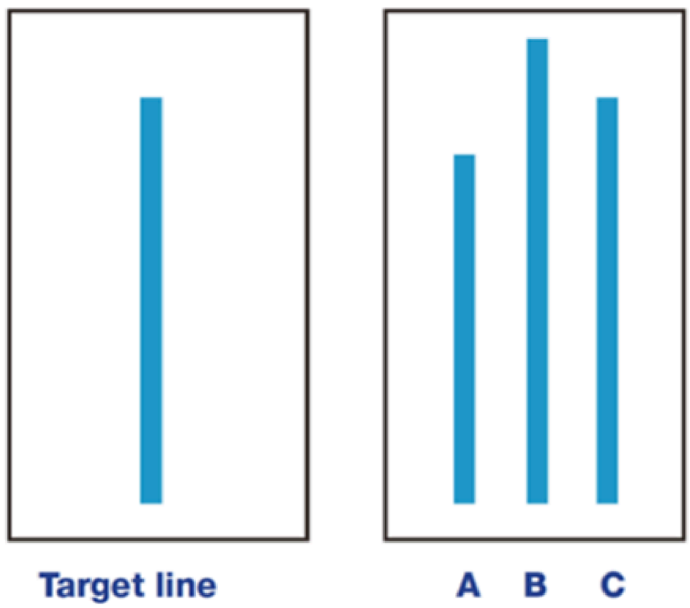
Rothbart et al. find that when under low-memory load and in the multiple-exposure condition, the typical subject's recollection of desirable stimulus persons in the group, as well as his or her judgment of group attractiveness, both remained constant as the proportion of presentation of the same desirable stimulus persons increased. This means that if a new individual joined the multiple-exposure group, say Fred, who demonstrated two instances of generosity, then this would not alter the typical subject's recollection and judgment of the group's desirability. On the contrary, when the proportion of different desirable stimulus persons increased in the single-exposure condition, the subject's judgment of group attractiveness increased proportionally (e.g., if sincere Sam was added to the single-exposure condition, the experiment's typical subject would increase his or her judgment of group attractiveness, and this increased amount would itself increase as additional desirable stimulus persons were added to the group). When under a high-memory load, the subject's recollection and judgment of the group's desirability increased proportionately as desirable stimulus persons were added under either the single- or multiple-exposure condition.

Therefore, the authors conclude that subjects under low-memory load organize their perceptions of group desirability around the preponderance of desirable stimulus persons in the group, while subjects under high-memory load organize their perceptions around the group as a whole regardless of whether an increase in desirability materializes through the addition of single encounters with desirable stimulus persons (single-exposure condition) or multiple encounters (multiple-exposure condition). In other words, the extent to which *Homo sapiens* stereotype depends upon how loaded their memories are with interactions with different members of the group being judged.[4]

CONFORMITY

In what has since been deemed a classic test of the social stigma of Conformity, Asch (1951) devised an experiment to test the extent to which a "naive subject" might choose to conform. In this experiment, the naive subject conforms when he selects an obviously wrong answer to a simple question as a result of having been unknowingly placed in a group of "stooges." Stooges had been pre-programmed to select the obviously wrong answer *en masse*.

The experiment presents the naive subject with the following figure.



"Psychology-asch-1951", by Saul McLeod, licensed under CC BY 3.0

He is then, eventually, asked which line in the box on the right—A, B, or C—is most like the Target line in the box on the left. Clearly, the correct answer is line C. However, in this experiment each of the stooges answers aloud and individually *before* the naive subject. In some treatment trials, each stooge answers line A, while in others they answer line B. The naive subject conforms if he answers line A in the former case and line B in the latter.

Asch (1951) conducted 18 trials total, 12 of which were treatment trials including stooges, and 6 of which were control trials without stooges (i.e., where no subjects in the group had been pre-programmed to say line A or line B). In the treatment trials, 75% of the naive participants conformed.

4. In a second experiment, Rothbart et al. test whether group members who are considered most salient have a disproportionate impact on our impressions of the group as a whole, in particular, whether "extreme individuals," by being novel, infrequent, or especially dramatic, are more available in memory, and thus, overestimated when judging their presence in the group. The authors ran the experiment with extreme instances of physical stimuli (men's heights) and social stimuli (unlawful behavior). In accord with the authors' predictions, subjects gave significantly higher estimates of the number of stimulus persons in the groups with extreme conditions—over six feet tall (in the physical stimuli treatment) and criminal acts (in the social stimuli treatment).

In the control groups, only 1% of subjects gave incorrect answers (i.e., 99% of the subjects chose line C).[5]

What do you think would have been the outcome of this experiment if the naive subjects in the treatment groups would have been from the *Homo economicus* species rather than *Homo sapiens*?

HINDSIGHT BIAS

Fischhoff and Beyth (1975) engaged students at Hebrew University and Ben Gurion University of the Negev in Israel in an interesting experiment regarding the students' predictions and recollections of then-President Nixon's diplomatic visits to China and the USSR in the first half of 1972. It turns out that 75% of the students recalled having assigned higher probabilities than they actually had to events that they believed had happened. Specifically, shortly after Nixon's visit to China or the USSR, these students erroneously believed that they had more accurately predicted before Nixon's visit that Nixon would make the trip. And 57% of the students erroneously recalled having assigned lower probabilities than they actually had to events they believed had not happened. Specifically, shortly after Nixon's visit to China or the USSR, these students erroneously believed that they had more accurately predicted before Nixon's visit that Nixon would not visit somewhere else, other than China or the USSR. Ouch!

These exaggerations of predictive accuracy are reflective of what is known as Hindsight Bias.[6]

LESS IS MORE

Redelmeier et al. (2003) report on a study of roughly 700 patients who underwent colonoscopies. The patients self-reported the intensity of pain on a 10-point scale (0 = "no pain at all" to 10 = "intolerable pain") every 60 seconds during the procedure. By random assignment, half of the patients' procedures lasted a relatively short amount of time. The other half had a short interval of time added to the end of their procedure during which the tip of the colonoscope remained in their rectums (egad!).

The experience of each patient varied considerably during the procedure. As an example, suppose Patient A's procedure lasted 14 minutes, while Patient B's lasted 29 minutes. Later, shortly after their respective procedures, we would expect Patient B to have reported incurring more pain overall (again on a 10-point scale) since the "integral of pain" is larger for Patient B. On the contrary, the researchers found that on average those patients who underwent the extended procedure (like Patient B) rated

5. For a nice synopsis of the original Asch conformity experiment, along with more recent conformity findings, see McLeod (2018).

6. Hindsight Bias is a special case of biases associated with "optimistic overconfidence," of which there are several examples. For instance, Svenson (1981) questioned drivers in the US and Sweden about their overall driving skills. 93% of US drivers and roughly 70% of Swedish drivers believed they drive better than the respective average drivers in their countries, indicating general overconfidence in their driving skills. In some instances, people may express Hindsight Bias in a less egotistical manner by believing that others possess the same level of skill or knowledge as they do. This is what's come to be known as the "curse of knowledge." Whereas Hindsight Bias results when *Homo sapiens* look backward in time, a similar bias results when we look forward in time (i.e., when we predict the future), called Projection Bias. Projection Bias occurs when we do not change how we value options (e.g., decision outcomes) over time. Thus, we tend to ignore the impacts of certain factors that have changed in the intervening time, which can later lead to regret at having made the decision we did. As we will see in Chapters 4 and 5, Hindsight and Projection Biases lead to what's known as the "time-inconsistency problem" for *Homo sapiens*, where what we believe we will want at some future time disagrees with what we actually want (and therefore choose) at that future time (which can be thought of as an inter-temporal version of a preference reversal).

their overall experience as less unpleasant than those who underwent the shorter procedure.[7] This result exemplifies what's known as Less is More.

FLAT-RATE BIAS

Many services can be purchased on a per-use or flat-rate basis (e.g., per-month, per-season, or per-year). For example, mass transit in most cities can be paid for per ride or via a monthly or pre-paid transit pass. Health clubs allow you to pay per visit or on a monthly, annual, or punch-pass basis. Living in Utah where downhill skiing is considered by many to be a must-do, I can purchase a daily lift ticket each time I arrive at the mountain or pre-purchase a season's pass which allows me unlimited visits during the ski season. It seems only logical that for purchases like these, people decide how many times they will use the service during a given year and then choose to purchase on a per-use or a flat-rate basis, whichever costs them less. But such is not always the case. *Homo sapiens* are prone to what's known as Flat-Rate Bias.

For example, Della Vigna and Malmendier (2006) analyzed data from three U.S. health clubs with information on both the contractual choice and the day-to-day attendance decisions of approximately 8,000 members over three years. Members who chose a contract with a flat monthly fee of $70 attended on average 4.3 times per month. They, therefore, paid an effective price per visit of more than $17 even though they could have paid only $10 per visit using a 10-visit punch pass. On average, these members forgo savings of roughly $600 over the course of their memberships. Further, members who chose the monthly contract were 17% more likely to stay enrolled beyond one year than users pre-committing for a single year. Flat-Rate Bias, therefore, has relatively costly consequences for these members.

Della Vigna and Malmendier mention the leading explanations for their findings: overconfidence about future self-control and resolve. Overconfident members overestimate their future attendance as well as their resolve to cancel automatically renewed contracts. This latter manifestation of overconfidence—a lack of resolve when it comes to cancelling a contract—is what's known as a time inconsistency problem (or a Projection Bias). Time inconsistency arises when you make a suboptimal choice in the moment that is inconsistent with how you envisioned making that choice at an earlier period in time (e.g., as part of a larger plan). For example, you may originally sign the contract with automatic renewal with the understanding that if later on you find yourself working out less than expected, you will cancel. But as "later on" becomes "today," you're just too busy or forgetful to follow through with the cancellation.[8]

DIVERSIFICATION BIAS

As Read and Loewenstein (1995) point out, the rational model of *Homo economicus*' choice behavior

7. Nevertheless, those patients undergoing the extended procedure reported having experienced less pain in the final moments of their procedures than those undergoing the shorter procedure.

8. In a related study of members of three health clubs in Colorado, Gourville and Sorman (1998) found evidence of what's known as "payment depreciation," whereby payment for club membership has a diminishing effect on members' use of the club as time goes on. Memberships in these clubs were purchased on an annual basis with payments made semi-annually from the time of enrollment. For instance, a member joining in January would pay in January and June each year. The authors found that no matter when the month of payment occurred, there was a substantial spike in attendance immediately following payment. Approximately 35% of the average member's attendance during any six-month window occurred in the month of payment. In contrast, roughly 10% of attendance occurred during the fourth or fifth month after payment.

suggests that it is better to make choices in combination rather than separately, in other words, to frame the choices broadly rather than narrowly. For example, for dinner tonight, you should choose the restaurant you will eat at based upon the expected quality of the entrée, and then also envision the wine you will choose based upon the entrée. And before going out, you should select a matching outfit. It would probably be best if you combine the restaurant, dinner, wine, and outfit into a single interrelated choice. The advantages of a combined choice stem from the complementarities and substitutability between the individual choices. Only when choices are made in tandem can such interactions be accommodated optimally.

Previous experiments conducted by Simonson (1990) and Simonson and Winer (1992) found that if consumers combine their purchases at a single, initial point in time (to simultaneously provide for current and future consumption), they will choose more diverse bundles (i.e., exhibit more variety-seeking behavior) than if they make purchases sequentially at the various points in time when the goods are actually to be consumed. *Homo sapiens* being *Homo sapiens*, the question naturally arises as to whether consumers are prone to over-compensate when it comes to adding variety to their bundles. When consumers plan for more variety (in the simultaneous-choice setting) than they will subsequently desire (in the sequential-choice setting), they exhibit a Diversification Bias. Read and Loewenstein conduct a series of simultaneous- and sequential-choice laboratory experiments to explain potential underlying causes of this bias.

The authors identify a host of possible explanations for Diversification Bias stemming from (1) time contraction, when consumers subjectively shrink the inter-consumption interval when making an initial simultaneous choice, thus exaggerating the impact of satiation on their preferences, and (2) choice bracketing, when a simultaneous choice is presented to consumers in the form of a package and the most straightforward choice heuristic is to diversify.[9] In the experiments, roughly 375 subjects (undergraduate economics students at Carnegie Mellon University) are randomly assigned to groups tasked with making either simultaneous or sequential choices across three different snacks from among the following six: Snickers bars, Oreo cookies, milk chocolate with almonds, tortilla chips, peanuts, and cheese-peanut butter crackers. Read and Loewenstein ultimately find evidence of both time contraction and choice bracketing as underlying reasons for Diversification Bias.

THE BIAS BLIND SPOT

We conclude this chapter with a simple question (and provide a generally accepted answer to it). To what extent do *Homo sapiens* perceive themselves as less guilty of biasedness in their own thinking than others are in theirs—others such as the "average American" or "average classmate"? Pronin et al. (2002) couch this question of "asymmetry in perceptions of bias" as an informal hypothesis:

> "People think, or simply assume without giving the matter any thought at all, that their own take on the world enjoys particular authenticity and will be shared by other openminded perceivers and seekers of truth. As a consequence, evidence that others do not share their views, affective reactions, priorities regarding social ills, and so forth prompts them to search for some explanation, and the explanation most often arrived at is that the other parties' views have been subject to some bias that keeps them from reacting as the situation demands. As a result of explaining such situations in terms of others' biases, while failing to recognize the role of similar

9. The authors also test several hypotheses about why *Homo sapiens* may choose greater diversity in a simultaneous-choice setting, but not as a result of Diversification Bias. Rather, the reasons stem from mispredictions of taste, risk aversion and uncertain preferences, and information acquisition about a larger variety of commodities.

> biases in shaping their own perceptions and reactions, individuals are likely to conclude that they are somehow less subject to biases than the people whom they observe and interact with in their everyday lives." (pp 369-370)

The authors label this asymmetry as a special case of "naive realism" called the Bias Blind Spot. To test their hypothesis, Pronin et al. surveyed a group of 24 Stanford students enrolled in an upper-level psychology class on their susceptibility to eight different biases compared with both the average American and their average classmate. The biases were associated with (1) self-serving attributions for success versus failure (Self-Serving), (2) reduction of cognitive dissonance after having voluntarily made a choice (Cognitive Dissonance), (3) the halo effect (recall our discussion in Chapter 1) (Halo Effect), (4) biased assimilation of new information (Biased Assimilation), (5) reactive devaluation of proposals made by one's negotiation counterparts (Reactive Devaluation), (6) perceptions of hostile media bias toward one's group or cause (Hostile Media), (7) the fundamental attribution error associated with blaming the victim (FAE), and (8) judgments about the greater good that are influenced by personal self-interest (Self Interest).[10]

The authors point out that while none of these particular biases had previously been discussed in the class, participants may have learned about some of them in other psychology courses. Further, the descriptions of the biases used the neutral terms "effect" or "tendency" rather than the nonneutral term "bias." Thirteen participants were asked first about their susceptibility to each of the eight biases ("To what extent do you believe that you show this effect or tendency?") and then about the susceptibility of the average American to each ("To what extent do you believe the average American shows this effect or tendency?") The remaining 11 students rated the average American's susceptibility before their own. Ratings were made on nine-point scales anchored at 1 ("not at all") and 9 ("strongly"), with the midpoint of 5 labeled "somewhat."

Pronin et al. found that for each bias the students on average rate their susceptibility less than what they perceive as the average American's. For some biases (e.g., Self-Serving, FAE, and the Halo Effect), the difference is relatively large, while for others (e.g., Reactive Devaluation and Cognitive Dissonance), the difference is relatively small.[11] Interestingly, the students also rated their parents as less susceptible to each bias than the average American. In a separate study with a different sample of students, Pronin et al. find that, although not as strong, these results extend to the average student in another seminar course—a comparison group that is less hypothetical and more relevant to the participating students than the average American.

Alas, it appears that not only are *Homo sapiens* susceptible to a host of biases and fallacies and effects, but they are also susceptible to projecting their biases onto others, creating a ripple effect throughout society. This leads Pronin et al. to conclude that,

> "In the best of all possible worlds, people would come to recognize their own biases and to recognize that they are no less susceptible to such biases than their adversaries. In the imperfect world in which we live, people should at least endeavor to practice a measure of attributional charity. They should assume that the "other side" is just as honest as they are (but not more honest) in describing their true sentiments—however much these may be distorted by defensiveness, self-interest, propaganda, or unique historical experience." (p 380)

10. These biases are measured by Pronin et al. in an objective reality context. Related research conducted by Cheeks et al. (2020) shows that people are similarly susceptible to the Bias Blind Spot in the subjective domain of art appreciation.

11. In Chapter 5 we will learn how researchers discern differences like these on a more formal, statistical basis.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013), and those marked with a "‡" are adopted from Cartwright (2014).

1. † *Homo economicus* prefers information that is accurate no matter how it relates to her current hypothesis. She continues to seek new information until she is certain enough of the answer that the cost of additional information is no longer justified by its degree of uncertainty. As we know, Confirmation Bias among *Homo sapiens* can lead to overconfidence, which impedes the individual from fully recognizing the level of uncertainty she faces. What implications are there for an information search by individuals displaying Confirmation Bias? When do these individuals cease to search for additional information? What might this imply about people who have chosen to cease their education efforts at various phases? How might education policy be adjusted to mitigate this bias among individuals who terminate their educations at different levels (e.g., before earning an Associate's or Bachelor's degree)?

2. Discuss how Status Quo Bias and the Mere Exposure Effect are related to each other.

3. Explain how Status Quo Bias has a negative impact on your life. Can you think of an example of how this bias impacts your life positively?

4. ‡ Why might the use of a heuristic stem from an underlying bias? Give an example of a heuristic presented in Chapter 1 that accompanies a bias presented in this chapter. Should we emphasize how clever people are for utilizing good heuristics or how deficient they are for being biased in the first place?

5. Give an argument for why people who believe in extrasensory perception (ESP) or are prone to superstition are more likely to exhibit Confirmation Bias than people who are not.

6. Why might people who are prone to Status Quo Bias also be prone to Confirmation Bias?

7. Explain how Confirmation Bias both differs from and is similar to Jumping to Conclusions.

8. As mentioned in the discussion of the Law of Small Numbers, basketball fans tend to believe they are witnessing a "hot hand" when a player makes a series of shots during a game. Gilovich, et al. (1985) dismissed the hot hand as a myth. Recall that based upon a season's worth of data, the authors found that if a given player on a given night had just missed one shot, then on average, he hit 54% of his subsequent shots. Likewise, if the player had just hit one shot, he hit 51% of his remaining shots. After hitting two shots, he then hit 50% of subsequent shots. The estimated correlation coefficient between the outcome of one shot and the next was a statistically insignificant –0.039, suggesting that shooting streaks are an illusion. Take a look at this YouTube video of Houston Rocket's star player Tracy McGrady's amazing performance in the final 35 seconds of a crucial game against the San Antonio Spurs

on December 9, 2004, a performance that many basketball fans believe cumulated into the greatest series of plays in National Basketball Association history. If this wasn't a hot hand, then what was it?

9. One way to demonstrate the implication of the Law of Small Numbers is to show how small samples are inclined to result in greater extremes in terms of deviations from expected lottery outcomes. To see this, suppose a jar is filled with a total of six marbles—three green and three red. You randomly choose three marbles without replacement. Calculate the probability of choosing the extreme of either three red or three green marbles. Now, assume the jar is filled with a total of eight marbles—four green and four red. You randomly choose four marbles without replacement. Show that the probability of choosing the extreme of either four red or four green marbles is now less than it was when the jar was filled with three green and three red marbles.

10. Name a bias in this chapter that is likely to occur as a result of having adopted the Affect Heuristic. Name a bias that is a likely result of having adopted the Availability Heuristic.

11. Logicallyfallacious.com poses two questions. Answer each one. What fallacy are these two questions setting you up to fall victim to? *Question 1*: While jogging around the neighborhood, are you more likely to get bitten by someone's pet dog, or by any member of the canine species? *Question 2*: Sarah is a thirty-something-year-old female who drives a mini-van, lives in the suburbs, and wears mom jeans. Is Sarah more likely to be a woman or a mom?

12. Can you think of a situation in your own life where undertaking a "premortem" would likely help you avoid falling victim to the Planning Fallacy?

13. In the Bayes Rule example of Stereotyping, how low would the probability that a Hello taxi was involved in the accident given the witness's testimony (i.e., $Pr(H|W)$) have to be in order for the probability of a Hello taxi having been in the accident to equal 15%?

14. Can you think of an example where "less is more"? Henry David Thoreau, the 19th Century American naturalist, essayist, and philosopher, did so when he wrote in *Life in the Woods* (a.k.a *Walden Pond*) that a man's wealth is determined by the number of things that he can live without.

15. Think of a way in which you have conformed to a social norm. Describe both the social norm and how you went about conforming to it. In what ways has your conformity both benefitted and harmed you?

16. When a bank recently stopped charging per transaction (e.g., check fees and fees per phone-banking transaction) and changed to a per-month flat rate, their revenues went up by 15%. What is the most likely cause of this response among the bank's customers?

Media Attributions

- Figure 7 (Chapter 2) © Saul McLeod is licensed under a CC BY (Attribution) license

PART II.

SECTION 2 - *HOMO ECONOMICUS* VERSUS *HOMO SAPIENS*

We now turn to a discussion of the explicit axioms that distinguish *Homo economicus*, and describe a central theory (known as expected utility theory) that evolved from these axioms in the mid-20th century. The first two axioms are known as the Principal Rationality Axioms, and the second set of five axioms are known as Additional Rationality Axioms.[1] While we assume that *Homo economicus* would never violate any of them, the extent to which *Homo sapiens* violate these axioms has been a foundational question in behavioral economics. In this section, we take a deep dive into the comparison between *Homo economicus* and *Homo sapiens*.

We begin in Chapter 3 with an introduction to the axioms and principles that distinguish the rational choice behavior of *Homo economicus.* We then examine the implications of these axioms and principles with respect to how economists have traditionally depicted *Homo economicus*' preferences, specifically their risk preferences. In Chapter 4, we explore how behavioral economists have adjusted the rational-choice model of *Homo economicus* to account for the reality of *Homo sapiens*' choice behavior. We introduce two new theories—Prospect and Regret Theories—that provide the two main frameworks within which these adjustments are characterized. In Chapter 5, we investigate a variety of laboratory experiments that have measured the extent to which *Homo sapiens* deviate from the rationality axioms of Chapter 3. Chapter 6 presents additional laboratory experiments designed to test the implications of the theories advanced in Chapter 4.

1. Mas-Colell et al. (1995) include two additional axioms among the Principal Rationality Axioms known as the Reflexivity and Continuity Axioms. Although important for understanding the preference-based approach to microeconomic theory, these two axioms are not as germane to the ensuing discussion. Therefore, we leave their understanding to the more-interested reader of microeconomic theory.

CHAPTER 3.

THE RATIONALITY OF *HOMO ECONOMICUS*

PRINCIPAL RATIONALITY AXIOMS*

COMPLETENESS AXIOM

Suppose *Homo economicus* faces two lotteries, which we will denote as lotteries L and L', where both lotteries are taken from what is known as the space of available lotteries $\mathcal{L}$. Then, it must be the case that either $L \succsim L'$, $L \precsim L'$, or $L \sim L'$. What the previous sentence says is that *Homo economicus* either likes lottery L at least as much as lottery L' (i.e., $L \succsim L'$), likes lottery L' at least as much as lottery L (i.e., $L \precsim L'$), or is indifferent between the two lotteries (i.e., $L \sim L'$). This is known as the Completeness Axiom. For future reference, we will use the equivalent terminology "weakly preferred to" rather than "likes at least as much" when referring to the preference relation $\succsim$.

TRANSITIVITY AXIOM

Given any third lottery L'' taken from the space of available lotteries $\mathcal{L}$, if $L \succsim L'$ and $L' \succsim L''$, then $L \succsim L''$. In other words, *Homo economicus* would never fall victim to a "preference reversal," whereby she makes choices that contradict her stated preference ranking. This is known as the Transitivity Axiom.

So that we are clear on what a lottery is, here is an example of three possible lotteries L, L', and L''.

L	L'	L''
60% chance to win \$200	50% chance to win \$100	55% chance to win \$20
40% chance to lose \$100	50% chance to lose \$70	45% chance to lose \$20

As a final note, it is worth pointing out that the Principal Rationality Axioms imply the existence of what's known as a utility function representing an individual's preferences over lottery space $\mathcal{L}$, specifically, $U : \mathcal{L} \to R$ such that $L \succsim L' \Longleftrightarrow U(L) \geq U(L')$. Let's unpack this mathematical statement. The first part of the statement (i.e., $U : \mathcal{L} \to R$) says that utility function U magically translates an individual's preferences for the different lotteries that make up lottery space $\mathcal{L}$ into real numbers. The real numbers, by the way, are measured in what's known as "utils," or units of happiness. For example, if $U(L) = 100.3$, then the individual facing lottery L gets 100.3 units of happiness just from the opportunity of being able to play the lottery.

The second part of the statement (i.e., $L \succsim L' \Longleftrightarrow U(L) \geq U(L')$) says that the statement "lottery L is weakly preferred to lottery L'" (i.e., $L \succsim L'$) is equivalent to the statement "the utility level obtained from lottery L is no less than the utility level obtained from lottery L'" (i.e., $U(L) \geq U(L')$).

ADDITIONAL RATIONALITY AXIOMS*

DOMINANCE AXIOM

If $L \succsim L'$ in all respects (i.e., L's expected win outcomes are larger than L''s, and L's expected loss outcomes are lower), then $L \succsim L'$. To see what we mean by "expected win" and "expected loss" outcomes, refer to the example lotteries above in the discussion of the Transitivity Axiom. Lottery L 's expected win and expected loss outcomes, respectively, are $120 (0.6 x $200) and $40 (0.4 x $100), while lottery L''s are $50 (0.5 x $100) and $35 (0.5 x $70). The final part of this axiom's statement, $L \succ L'$, states that lottery L is "strictly" preferred to lottery L'. An equivalent way to say "strictly preferred to" is to say "likes more than." This is known as the Dominance Axiom.

Again referring to the example above, is $L \succ L'$ by the Dominance Axiom? The answer is "no." Although lottery L's expected win of $120 is larger than lottery L''s expected win of $50, L's expected loss of $40 is also larger than L''s expected loss of $35. If lottery L's expected loss had instead been some amount less than $35, we could have concluded that $L \succ L'$ by the Dominance Axiom.

INVARIANCE AXIOM

If an individual's preference ordering of different lotteries does not depend on how the lotteries are described, then the individual's preferences satisfies the Invariance Axiom. We do not have a mathematical expression for this axiom. But there will be plenty of examples as we explore the seminal experiments conducted by Kahneman and Tversky, experiments that motivated their famous theories in behavioral economics. We will be learning about these theories later in this section of the textbook.

SURE-THING PRINCIPLE

If $L \succsim L'$ in any possible known state of the world, then $L \succsim L'$ even when the state of the world is unknown. This is known as the Sure-Thing Principle. Similar to the Invariance Axiom, we do not have a mathematical expression to contend with. We will soon see an example.

The Sure-Thing Principle implies that an individual does not need to consider uncertainties when making a decision if the individual deems the uncertainties to be irrelevant. For example, if an investor has decided to purchase a company's stock regardless of its upcoming earnings report, then it makes no sense for the investor to worry about whether the company will ultimately report a profit or a loss. His preference for the company's stock is unaffected by the uncertainty surrounding its reported profit level.

INDEPENDENCE AXIOM

An individual's preferences defined over lottery space $\mathcal{L}$ satisfies the Independence Axiom if for any three lotteries L, L', and L'', and some constant α that lies between 0 and 1 (non-inclusive of 0 and 1), we have,

$$L \succsim L' \Longleftrightarrow \alpha L + (1-\alpha) L'' \succsim \alpha L' + (1-\alpha) L''.$$

This last part of the axiom is a bit ugly. So let's look at an example.

Let $\alpha = 0.6$ and recall L, L', and L'' from the previous example:

$\underline{L}$	$\underline{L'}$	$\underline{L''}$
60% chance to win \$200	50% chance to win \$100	55% chance to win \$20
40% chance to lose \$100	50% chance to lose \$70	45% chance to lose \$20

$\underline{\alpha L + (1-\alpha)L''}$	$\underline{\alpha L' + (1-\alpha)L''}$
36% chance to win \$200	30% chance to win \$100
22% chance to win \$20	22% chance to win \$20
24% chance to lose \$100	30% chance to lose \$70
18% chance to lose \$20	18% chance to lose \$20

Thus, in adherence to the Independence Axiom, if, say, an individual weakly prefers lottery L to lottery L', then that individual will also weakly prefer lottery $\alpha L + (1-\alpha)\, L''$ to lottery $\alpha L' + (1-\alpha)\, L''$. By the looks of our example, the comparison between lottery $\alpha L + (1-\alpha)\, L''$ and lottery $\alpha L' + (1-\alpha)\, L''$ is likely to be difficult for *Homo sapiens*, who would thus be prone to violate this axiom. Hats off to *Homo economicus* for not violating the Independence Axiom.

SUBSTITUTION AXIOM

An individual's preferences defined over lottery space $\mathcal{L}$ satisfies the Substitution Axiom if, for any two lotteries L and L', and again some constant α that lies between 0 and 1 (non-inclusive of 0 and 1), we have, $L \succsim L' \iff \alpha L \succsim \alpha L'$. This is a simpler axiom than the Independence Axiom, but let's look at an example anyway.

Let $\alpha = 0.6$ and recall earlier L and L' from the previous examples:

$\underline{L}$	$\underline{L'}$
60% chance to win \$200	50% chance to win \$100
40% chance to lose \$100	50% chance to lose \$70

Then,

$\underline{\alpha L}$	$\underline{\alpha L'}$
36% chance to win \$200	30% chance to win \$100
24% chance to lose \$100	30% chance to lose \$70

In adherence to the Substitution Axiom, if, say, an individual weakly prefers lottery L to lottery L', then that individual will also weakly prefer lottery αL to lottery $\alpha L'$. By the looks of our example, *Homo sapiens* should at least be able to make the necessary comparisons between each of these lotteries and thus potentially satisfy this axiom.

HOMO ECONOMICUS AND THE EXPECTED UTILITY FORM**

One of the key assumptions of rational choice theory is that *Homo economicus* maximizes expected utility based upon his or her "total wealth" associated with the different outcomes of a lottery. It turns out that because *Homo economicus* satisfies the Independence Axiom, his or her expected utility is expressed in what is known as the "expected utility form" (EUF). Mathematically, we say that utility function $U : \mathcal{L} \rightarrow R$ has the EUF if there is an assignment of utility and probability values ($u(w_i)$ and π_i, respectively) defined over a given lottery's outcomes $i = 1, \ldots, N$, such that $U(L) = \sum_i (\pi_i \cdot u(w_i))$, where w_i represents the individual's total wealth associated with outcome i, i.e., the sum of his or her initial wealth level plus (or minus) the winnings (or losses) associated with outcome i.

Ouch. We had better jump to an example where, for simplicity, we set N = 2 (i.e., we consider lotteries with only two possible outcomes—a win and a loss), like the lotteries L, L', and L'' presented in the previous examples. Thus, in this case expected utility $\sum_i (\pi_i \cdot u(w_i))$ can be written as $\pi_1 \cdot u(w_1) + \pi_2 \cdot u(w_2)$.

Suppose an individual has initial wealth of $100 and his or her utility function $u(w_i) = \sqrt{w_i}$. Then, using the values from the previous examples,

$$U(L) = 0.6\sqrt{300} + 0.4\sqrt{0} = 10.4$$
$$U(L') = 0.5\sqrt{200} + 0.5\sqrt{30} = 9.8$$
$$U(L'') = 0.55\sqrt{120} + 0.45\sqrt{80} = 9.5$$

Thus, $L \succ L' \succ L''$

You are probably wondering where all of the numbers in the example are coming from. Let's first take a look at the value for $U(L)$, the expected utility associated with lottery L. Starting with the values $\sqrt{300}$ and $\sqrt{0}$, note that according to lottery L, if the individual wins, then he wins $200, which, added to his initial wealth of $100, results in $300. Similarly, if the individual loses, then he loses $100, which, subtracted from his initial wealth of $100, results in $0. The square roots of 300 and 0 come about because the individual's utility function for this example is specified as $u(w_i) = \sqrt{w_i}$, where, in this case, i = 1 corresponds to a win outcome of $200, and i = 2 corresponds to a loss outcome of $100.

Lastly, the values π_1 = 0.6 and π_2 = 0.4, respectively, correspond to the 60% probability of the win outcome occurring with lottery L, and the 40% probability of the loss outcome occurring. Thus, via the expected utility form, we have

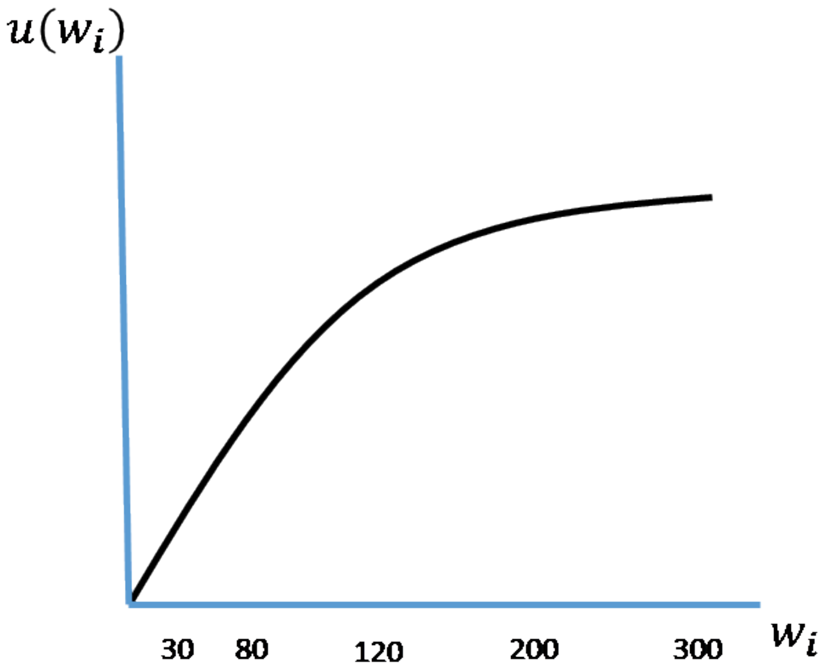
$U(L) = \pi_1 \cdot u(w_1) + \pi_2 \cdot u(w_2) = 0.6\sqrt{300} + 0.4\sqrt{0}$ = 10.4 utils.

The expected utility values for lotteries L' and L'' are derived in the same manner. Since $U(L) > U(L') > U(L'')$ in this example, the result $L \succ L' \succ L''$ naturally follows.

The utility expression $u(w_i) = \sqrt{w_i}$ indicates that our individual exhibits "risk aversion" with respect to total wealth level w_i. This is due to the utility expression itself exhibiting diminishing marginal utility in w_i. Thus, diminishing marginal utility in w_i is synonymous with risk aversion. To further understand this relationship between diminishing marginal utility and risk aversion, we turn to a graphical analysis.

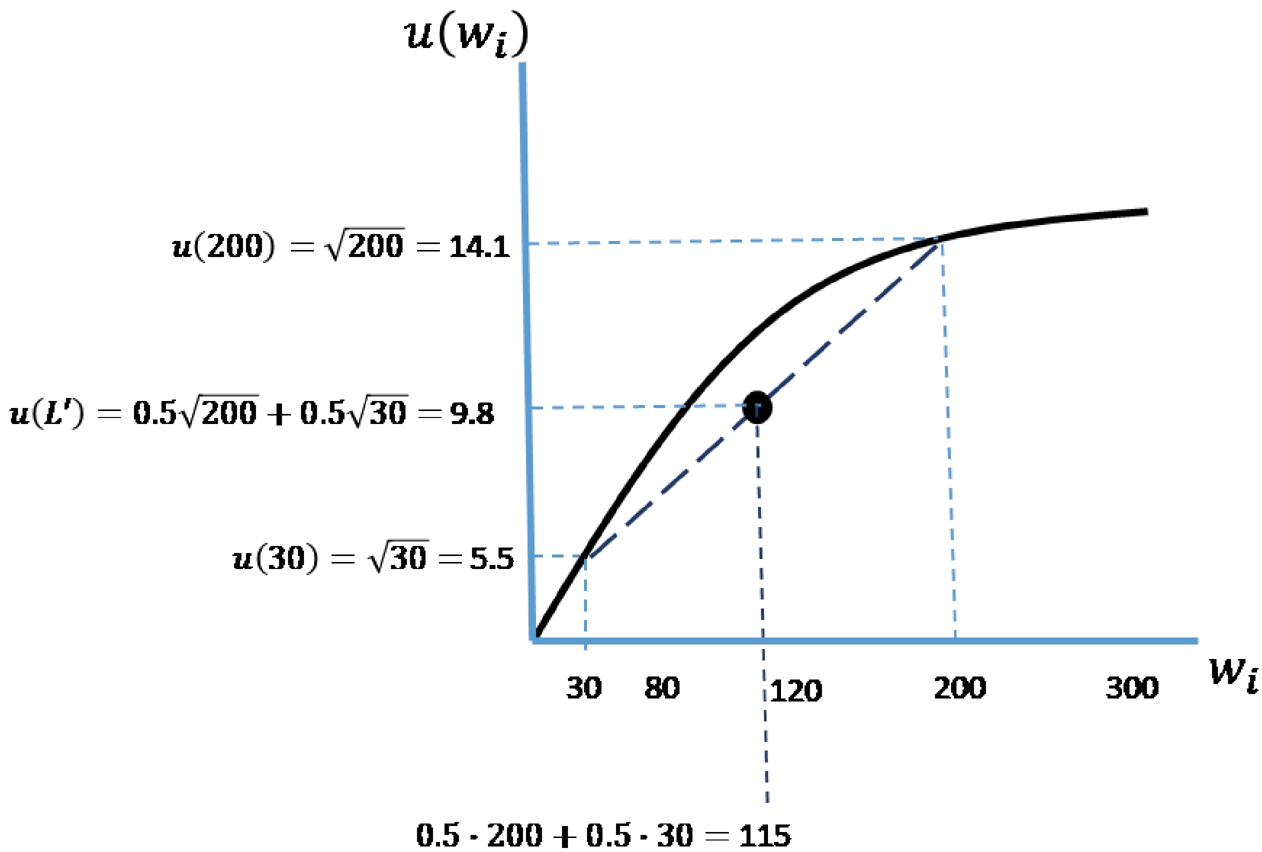
In Figure 3.1 below, we depict $u(w_i) = \sqrt{w_i}$ defined on a continuum from the lowest total wealth of $0 at the graph's origin to the largest total wealth of $300. We also indicate the other total wealth levels of $30, $80, $120, and $200 associated with lotteries L' and L''. Note that $u(w_i) = \sqrt{w_i}$ indeed exhibits diminishing marginal utility in w_i, or to state it yet another way, $u(w_i) = \sqrt{w_i}$ is "concave" in w_i.

Figure 3.1. Utility Function Defined Over Wealth Levels



For sake of example, let's now superimpose the values associated with lottery L' on this graph, as shown in Figure 3.2.

Figure 3.2. Utility Function Defined Over Wealth Levels with Superimposed Lottery Values.

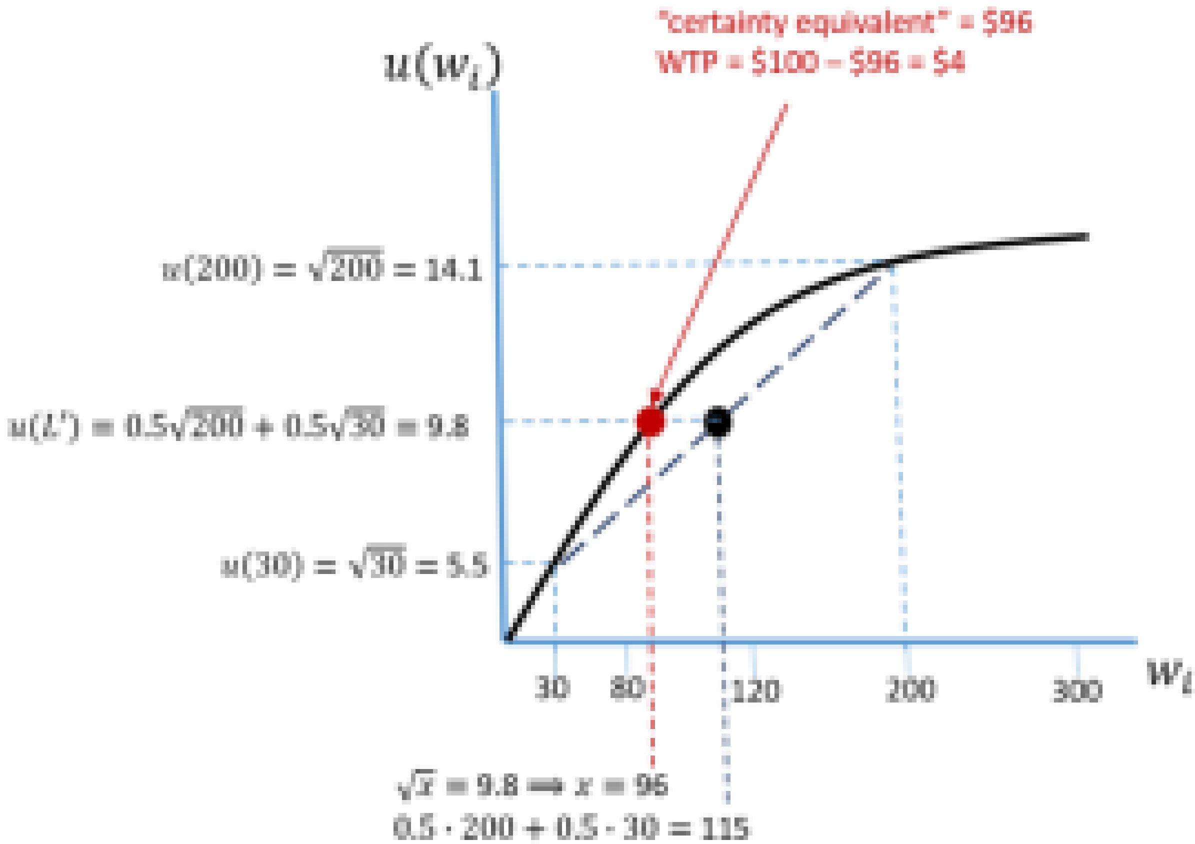


Begin by noting that with lottery L' the individual's total wealth when she wins equals \$200 (which, as indicated on the graph's vertical axis, corresponds to $u\left(200\right) = \sqrt{200}$ = 14.1 utils), and equals \$30 when she loses (corresponding to $u\left(30\right) = \sqrt{30}$ = 5.5 utils). We already know from the information provided in the example that $U\left(L'\right) = 0.5\sqrt{200} + 0.5\sqrt{30}$ = 9.8 utils, as indicated on the graph's vertical axis. What the example did not tell us is that the midpoint on the line segment connecting the individual's utility values at lottery L''s loss outcome of \$30 and win outcome of \$200 corresponds to (0.5 x \$200) + (0.5 x \$30) = \$115.[1]

We can identify two values in this graph that provide quantitative measures of an individual's risk aversion. As shown in Figure 3.3, the measures are known as "certainty equivalent" and "willingness-to-pay" (WTP) to avoid having to play the lottery in the first place.

Figure 3.3. Certainty Equivalent and WTP

1. Note that our reference point for this example—the midpoint of the line segment—corresponds to lottery L''s 50%-50% split in probability of a win and a loss. If, for example, the lottery's split had instead been 60%-40% in favor of winning, then the reference point would be located further to the northeast on the line segment, corresponding to a w_i of 0.6 x (\$200 + \$30) = \$138.



As depicted in the graph, the individual's certainty equivalence is found by drawing a line horizontally from the midpoint of the line segment to its intersection with the utility function. Mathematically, this intersection corresponds to the value of w_i (in the graph denoted as x) that solves for the expected utility value of lottery L', 9.8 utils. This value is $96, which, because the individual in this example is risk averse, is less than $115.[2] The individual's WTP is then calculated as $100 – $96 = $4. In other words, the individual is willing to pay $4 (out of his initial wealth of $100) to avoid having to play lottery L' in the first place.

Before leaving this discussion of the expected utility form, consider the following thought experiment:

> Choose between lotteries **A** and **B**:
>
> **A** 80% chance to win $4,000
>
> **B** win $3,000 for certain

If you chose lottery B, then you are considered risk averse. That's because the expected value of playing lottery A (0.8 x $4,000 = $3,200) is larger than the certain value of lottery B, $3,000. If you weren't averse to risk you would have chosen lottery A.

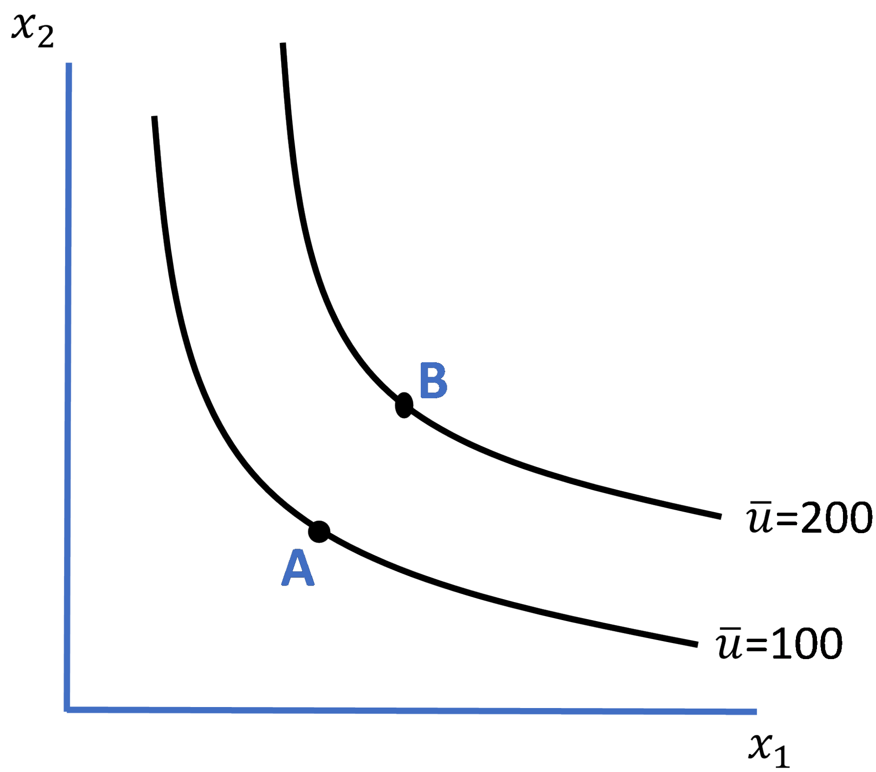
2. By contrast, if the individual had been assumed "risk neutral," in which case her utility function would be drawn linear with respect to wealth w_i, certainty equivalent would per force be equal to $115. What would be the result if our individual was instead "risk seeking," also known as "risk loving"?

Because we are precluded from ever deriving your utility function, we cannot calculate your certainty equivalence, *per se*. So, this type of thought experiment proffers an admittedly expedient way to gauge whether you are risk averse. Can you think of an alternative experiment that would enable us to actually obtain your certainty equivalence?[3]

HOMO ECONOMICUS AND THE INDIFFERENCE CURVE**

Before concluding this chapter on the rationality of *Homo economicus*, we introduce another important concept in Figure 3.4—the "indifference curve." In doing so, we depart from the world of expected utility defined over *Homo economicus*' possible wealth levels and venture into a world where *Homo economicus* chooses different levels of actual commodities to consume. In other words, we model *Homo economicus*' choices in the context of a marketplace rather than the context of a lottery.

Figure 3.4. *Homo economicus*' Indifference Curves



In Figure 3.4, let variables x_1 and x_2 represent the physical amounts of any two commodities 1 and 2 that *Homo economicus* might choose to consume (as a bundle) at some given point in time, and constant $\bar{u}$ represent some predetermined utility level, say 100 utils. An indifference curve defined for $\bar{u}$ = 100 depicts the tradeoffs *Homo economicus* is willing to make between the respective amounts of

3. Suppose I would have confronted you with the following lottery: 80% chance of winning \$4,000 and 20% chance of losing \$3,000. I then ask you two questions: (1) what is your initial wealth, and (2) what are you willing to pay to avoid having to play this lottery (WTP)? Suppose you answer that your initial wealth is \$10,000 and your WTP is \$100. Recalling our previous graphical analysis, we could stop right here. Because your WTP is greater than zero, we know that for this particular lottery, you are risk averse. But we can also go a bit further and calculate your corresponding certainty equivalence for this lottery. From our graphical analysis we learned that certainty equivalence is the difference between your expected total wealth from playing the lottery and your WTP. Given that your initial wealth is \$10,000, and you have an 80% chance of winning \$4,000 and a 20% chance of losing \$3,000, your total expected wealth from playing the lottery is (0.8 x \$14,000) + 0.2 x \$7,000) = \$12,600. Subtracting your WTP of \$100 from \$12,600 results in a certainty equivalence of \$12,500.

commodities 1 and 2 that still result in 100 utils of utility.[4] Or, alternatively stated, the indifference curve depicts all of the bundles that yield *Homo economicus* the same utility level $\bar{u}$ = 100. Thus, *Homo economicus* would feel indifferent about owning any of the bundles on the indifference curve. Hence its name.

For those of you with a background in microeconomic theory, you will recall the stylistic version of *Homo economicus*' indifference curves as shown in Figure 3.4. By "stylistic" we mean that the curves are everywhere downward-sloping and convex (i.e., bowed-out) from the graph's origin. In this graph, we have drawn two indifference curves for the same individual—one corresponding to a utility level of 100 utils, the other to a utility level of 200 utils. It is no coincidence that the indifference curve associated with $\bar{u}$ = 200 lies everywhere above (to the northeast of) the indifference curve associated with $\bar{u}$ = 100. It is also no coincidence that the two curves are "well-behaved" (i.e., they are parallel and do not intersect each other).[5]

The reason why the indifference curve lying to the northeast is associated with larger utility is because of what is known as monotonicity, specifically the assumption *cum* property that an individual obtains more utility by consuming larger amounts of the commodities. Recalling that a given indifference curve depicts all of the different bundles of goods (in our case bundles of goods 1 and 2) that provide an individual with the same level of utility, if we pick any bundle on curve $\bar{u}$ = 100 (say, bundle A), we can always find a bundle on curve $\bar{u}$ = 200 (say, bundle B) that includes more of both goods and thus, by the Monotonicity Property, implies that utility is indeed higher on $\bar{u}$ = 200.

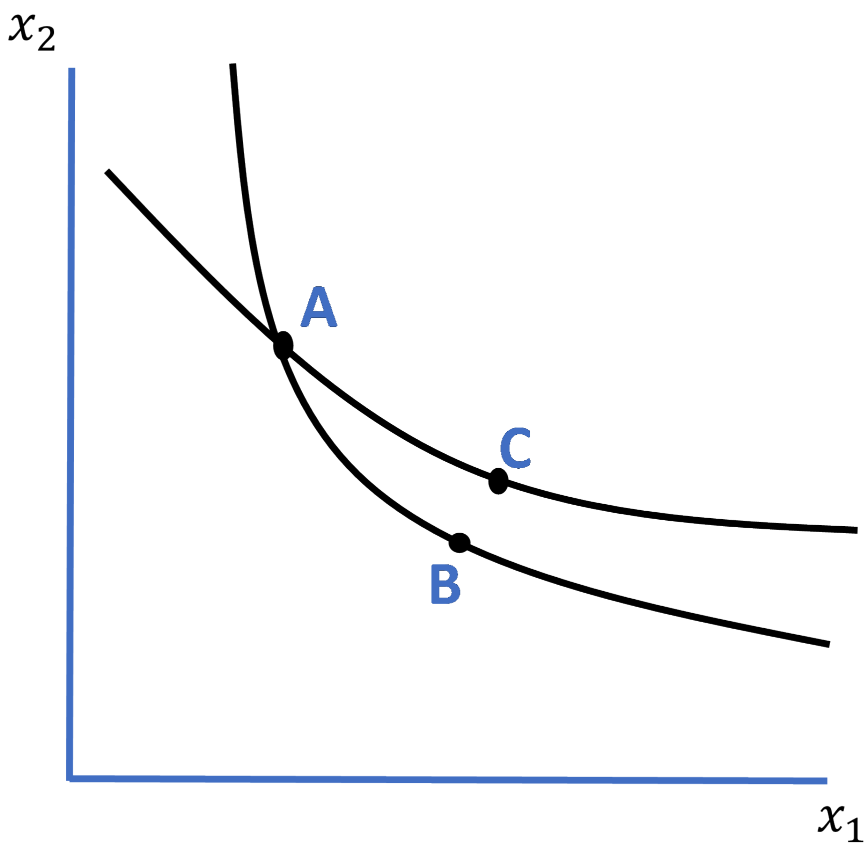
The fact that *Homo economicus*' indifference curves for $\bar{u}$ = 100 and $\bar{u}$ = 200 in Figure 3.4 do not intersect bears further mention. To see why, we need to recast our previous definition of the Transitivity Axiom we learned about in the context of lotteries to the context of commodities 1 and 2. As a reference point for the Transitivity Axiom, recall the earlier definition provided in the context of lotteries: if $L \succsim L'$ and $L' \succsim L''$, then $L \succsim L''$. In other words, if an individual weakly prefers lottery L to lottery L', and also weakly prefers lottery L' to lottery L'', then it must also be the case that the individual weakly prefers lottery L' to lottery L''. In the case of actual commodities, this definition of the Transitivity Axiom changes as follows: if an individual weakly prefers commodity bundle A to bundle B, and also weakly prefers bundle B to bundle C, then it must be the case that the individual weakly prefers bundle A to bundle C as well. See the parallel between these two contexts?

Now, assuming both the Transitivity Axiom and Monotonicity Property hold—as they do for *Homo economicus*—we can prove by contradiction that $\bar{u}$ = 100 and $\bar{u}$ = 200 do not intersect. Or, to state it another way, we can show that if the two curves do intersect, then transitivity and monotonicity cannot hold simultaneously. To see this, consider Figure 3.5 below, which includes two purposefully unlabeled indifference curves and three different consumption bundles A, B, and C.

Figure 3.5. A Violation of the Transitivity Axiom

4. For those of you who have seen indifference curves before, recall that the slope of the indifference curve at any given bundle of commodities x_1 and x_2 is the curve's marginal rate of substitution (MRS), which can be shown to equal the negative ratio of the individual's marginal utilities at that bundle. The MRS, then, is a marginal or continuous measure of the rate at which the individual is willing to tradeoff commodity 2 for more of 1.

5. The shape and location of the indifference curve is directly related to the individual's utility function as it is defined over x_1 and x_2. For example, it can be shown that indifference curves such as those depicted in the figure can be derived from what's known as a Cobb-Douglas utility function, where $u(x_1, x_2) = \sqrt{x_1 x_2}$.



Begin by noting that, via the definition of an indifference curve itself, because bundles A and B lie on the same indifference curve, the individual prefers the two bundles equally (i.e., obtains the same level of utility from each bundle). Similarly, because bundles A and C lie on the same indifference curve, the individual also values these two bundles equally. Now, according to the Transitivity Axiom, it must be the case that the individual prefers bundles B and C equally as well. But wait, by the Monotonicity Property, bundle C lies to the northeast of bundle B and therefore has more of both goods 1 and 2 in its bundle. So, this property tells us that the individual instead prefers bundle C over bundle B, which contradicts the earlier conclusion reached by the Transitivity Axiom. Hmmm. We have a contradiction here, one which cannot stand. Thus, the indifference curves cannot intersect each other when the Transitivity Axiom and Monotonicity Property hold simultaneously. Alternatively stated, the indifference curves for *Homo economicus* cannot intersect.

HOMO ECONOMICUS AND INTERTEMPORAL CHOICE***

The previous analyses of *Homo economicus*' preferences and choice behavior have been strictly "static" in the sense that we have restricted our representative individual to making decisions during a single time period: the here-and-now. But decisions are typically not made in such a vacuum time-wise. Instead, we (both *Homo economicus* and *Homo sapiens*) consider how decisions made today will affect future decisions. In other words, we take into account how money spent today (e.g., for a new shirt) will affect our ability to purchase something else in the future (e.g., the latest book about behavioral economics). Choices that span multiple time periods—not just the here-and-now, but also the 'there-and-later'—are known as "intertemporal," and the analysis of these choices is "dynamic" rather than static.

Certainly, the span of time we could conceivably account for in our dynamic analysis of *Homo* *economicus*' intertemporal choices is long, an entire lifetime. But to keep things tractable, and yet enable key comparisons between the intertemporal choice behavior of *Homo economicus* and (later, in Chapter 6) *Homo sapiens*, we make some simplifying assumptions. First, the time span under consideration is three periods rather than what could instead be considered a lifetime, or effectively an infinite number of periods.[6] Second, we assume preferences are stable, in particular, that the individual's utility function does not change shape (i.e., functional form) from one period to the next. Lastly, we assume that the individual's annual income level remains constant over time, as do relative prices. As a result, our analysis loses no generality by assuming that all prices are normalized to a value of one.

These last two assumptions imply that *Homo economicus* exhibits perfect foresight, which should come as no surprise. Perfect foresight means that *Homo economicus* can accurately predict how her preferences, annual income, or prices of the different commodities will change over time, and thus account for these changes in her decision problem at the outset. Therefore, *Homo economicus* cannot be 'tricked' into making what appear to be inconsistent choices over time.

To see this, recall our earlier definition of an individual's utility function, which allows us to model *Homo economicus'* three-period intertemporal decision problem as follows,

$$\underset{\{x_1, x_2, x_3\}}{Max} \quad u(x_1) + \delta u(x_2) + \delta^2 u(x_3)$$

Subject to,

$$W = (p_1 \cdot x_1) + (p_2 \cdot x_2) + (p_3 \cdot x_3)$$

where x_1, x_2, and x_3 represent the amount of a composite consumption good consumed by the individual in each period 1, 2, and 3, respectively; function $u(\cdot)$ represents the individual's utility defined over a given period's consumption level; parameter $0 < \delta \leq 1$ represents the individual's "discount factor," indicating the extent to which he is impatient for present, as opposed to future, consumption; prices p_1, p_2, and p_3 are the given, constant prices for periods 1, 2, and 3, respectively; and aggregate income $W = 3w$, where $w = w_1 = w_2 = w_3$ (the string of equalities imposes the previously mentioned constant period-by-period income level).

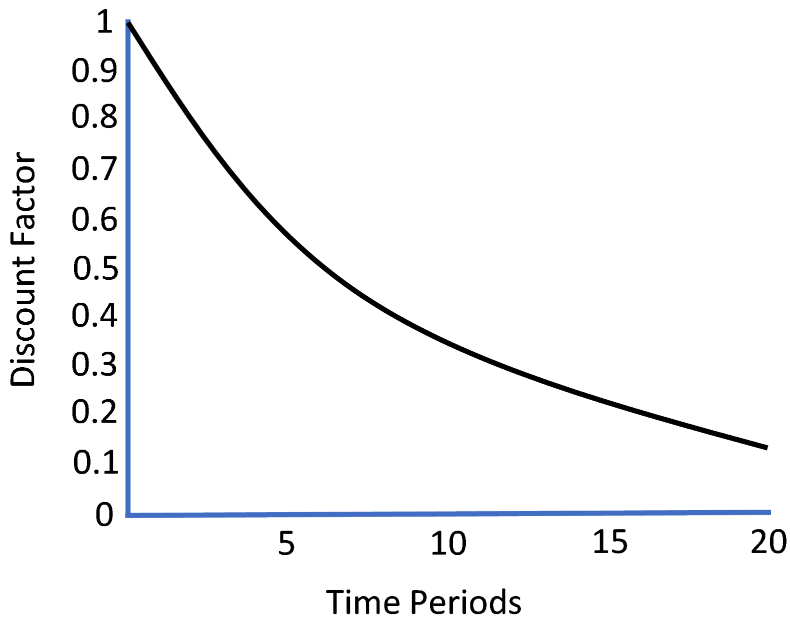
As with the utility function defined earlier over an individual's consumption bundle, the function $u(\cdot)$ is assumed to be increasing and concave in its respective period-by-period consumption levels. An impatient *Homo economicus* is represented by $\delta < 1$ (yes, as with risk aversion, rational intertemporal choice theory permits *Homo economicus* to be an impatient consumer). As such, in solving her intertemporal decision problem (i.e., choosing at the outset x_1, x_2, and x_3 to maximize aggregate utility across the three periods $u(x_1) + \delta u(x_2) + \delta^2 u(x_3)$), *Homo economicus* discounts her second-period utility level more than her first-period's, and in turn discounts her third-period utility more than her second period's. Such is the nature of impatience in the context of an intertemporal choice problem.

Because discount factor δ is constant across time but is nevertheless reduced in value through time exponentially (e.g., if there were a fourth period in our model, discounted utility would enter the individual's aggregate utility as "$+\delta^3 u(x_4)$ " and so on for additional periods), this pattern of discounting future utility is known as "exponential time discounting." As Figure 3.6 shows below (for $\delta < 1$), the schedule of 'discounted' discount factors over time charts out a negative exponential curve that is convex to the graph's origin and asymptotically zero with respect to the passage of time. In

6. Technically speaking, only two time periods are necessary to conduct dynamic analysis (e.g., today and tomorrow, or this year and next year). Hence, we could claim that our three-year time span extends beyond what is necessary here.

other words, as time stretches out further into the future, future utility is discounted progressively to a point where *Homo economicus* severely minimizes the influence of additions to his aggregate utility obtained in the distant future.

Figure 3.6. Exponential Time Discounting



Three particular features of *Homo economicus*' exponential discounting problem bear mentioning. First, in solving her intertemporal choice problem stated above, *Homo economicus* will optimally choose x_1, x_2, and x_3 such that her discounted marginal utility levels are equated across time. Let these optimal levels of consumption be denoted x_1^*, x_2^*, and x_3^*, respectively. This result ensures a choice profile where $x_1^* > x_2^* > x_3^*$, which is what we would expect the consumption profile of an impatient consumer to look like.

For those of you familiar with calculus, in particular solving constrained optimization problems, you will note that you can write this problem in its Lagrangian form as

$$\mathcal{L} = u(x_1) + \delta u(x_2) + \delta^2 u(x_3) + \lambda [W - x_1 - x_2 - x_3],$$

where $\lambda > 0$ represents the problem's Lagrangian multiplier, and for simplicity, we have normalized all prices to one (i.e., $p_1 = p_2 = p_3 = 1$). Obtaining the associated system of first-order conditions for this problem results in

$$u'(x_1) = \delta u'(x_2) = \delta^2 u'(x_3) = \lambda,$$

where $u'(\cdot)$ denotes the individual's marginal utility function. This string of equalities indicates that discounted marginal utility levels are equated across time. Given our underlying assumptions of $\delta < 1$ and diminishing marginal utility, the string of equalities in turn implies $x_1^* > x_2^* > x_3^*$.[7]

Second, *Homo economicus*' choice profile abides by what's known as "stationarity." In other words, her preferences for any increments to a given level of consumption in two different time periods depends upon the interval of time that passes between the two time periods (e.g., between periods 1

7. Since $x_1^* > x_2^* > x_3^*$ in the face of the underlying assumption that $w_1 = w_2 = w_3$, we effectively assume the individual borrows against future income (without interest penalty) to obtain the larger consumption level in period 1. Because the roles of borrowing and saving are extraneous to the insights we seek from the individual's choice problem, we lose nothing by casting the borrowing and savings decisions as implicit.

and 2), and not the specific points in time when the two respective increments could conceivably be consumed (e.g., periods 2 and 3).

For example, suppose *Homo economicus* chooses to consume the same base amount x in each period, and let the two different increments to base consumption be denoted as i_1 and i_2. Stationarity implies that if $u(x+i_1) > \delta u(x+i_2)$, where $u(x+i_1)$ represents the corresponding utility level in period 1 and $\delta u(x+i_2)$ represents discounted utility level in period 2, then $\delta u(x+i_1) > \delta^2 u(x+i_2)$ by exactly the same amount as $u(x+i_1) > \delta u(x+i_2)$, where $\delta u(x+i_1)$ represents a discounted utility level in period 2 and $\delta^2 u(x+i_2)$ represents a discounted utility level in period 3. This result is easy to see since $\delta u(x+i_1) > \delta^2 u(x+i_2)$ for periods 2 and 3 reduces to $u(x+i_1) > \delta u(x+i_2)$ for periods 1 and 2 after cancelling δ from each side of the former inequality. Since the two inequalities are identical through time, *Homo economicus*' preferences are stationary through time.

Note that the time intervals under comparison must be of equal length for stationarity to be assessed. For example, if $u(x+i_1) > \delta u(x+i_2)$ for consecutive periods 1 and 2, this does not necessarily imply that $u(x+i_1) > \delta^2 u(x+i_2)$ by the same amount across non-consecutive periods 1 and 3 (convince yourself of this fact). To further test your understanding of stationarity, suppose there was a fourth period of consumption. You should be comfortable seeing that if $u(x+i_1) > \delta^2 u(x+i_2)$ between periods 1 and 3, then $\delta u(x+i_1) > \delta^3 u(x+i_2)$ between periods 2 and 4 by the same amount.

Third, *Homo economicus*' preferences are "time consistent," In other words, given that he has chosen $x_1^*,\ x_2^*$, and x_3^* at the outset for any given set of δ, W, and prices p_1, p_2, and p_3, then if after having consumed at level x_1^* in period 1, the individual decides to re-solve his decision problem from that point forward (i.e., now starting in period 2), he will not deviate from choosing x_2^* and x_3^*. In this case, the individual effectively solves the two-period problem,

$$\underset{\{x_2, x_3\}}{Max} \quad u(x_2) + \delta u(x_3)$$

subject to,

$$W = 2w = (p_2 \cdot x_2) + (p_3 \cdot x_3),$$

which results in the same x_2^* and x_3^* as before. To see this result, first note that we can pull $\delta u'(x_2) = \delta^2 u'(x_3)$ from the string of three equalities derived from the individual's original decision problem (i.e., $u'(x_1) = \delta u'(x_2) = \delta^2 u'(x_3)$) and then cancel δ from each side of the equality, resulting in $u'(x_2) = \delta u'(x_3)$. But $u'(x_2) = \delta u'(x_3)$ is precisely the equality that results from the first-order conditions for this two-period problem. Given that nothing else has changed in this problem (i.e., the values for δ, w, and p_1, p_2, and p_3 are the same, as is the functional form of $u(\cdot)$ and the fact that x_1^* has already been chosen), no values for x_2 and x_3, respectively, solve $u'(x_2) = \delta u'(x_3)$ except x_2^* and x_3^*.

And if after having consumed at levels x_1^* and x_2^* for the first two periods, he then decides to re-solve his decision problem from that point forward (i.e., now starting in period 3), he does not have a problem left to solve—he has effectively locked himself into consuming x_3^* at that point. Therefore, because he chooses not to alter his consumption profile over time by having embarked on these period-by-period decision problems, we say that his choices are time consistent.

KEY TAKEAWAYS ON *HOMO ECONOMICUS*

So, what have we learned about *Homo economicus* thus far? For starters, we have learned that she is inviolable when it comes to making the miscalculations and misjudgments, exhibiting the biases, and falling victim to the fallacies and effects that bedevil *Homo sapiens*. Furthermore, *Homo economicus* is inextricably bound by the rationality axioms of Completeness, Transitivity, Dominance, Invariance, Independence, Substitutability, and the Sure-Thing Principle. And when confronted with the uncertainty of having to play a lottery, *Homo economicus* chooses to maximize expected utility derived over her total wealth. In the face of uncertainty, *Homo economicus* can be risk averse. And when it comes to a deterministic setting, such as when *Homo economicus* enters the marketplace to purchase actual commodities, the family of indifference curves that represent her preferences for different bundles of commodities are not permitted to intersect with each other.

In other words, *Homo economicus* is—to borrow the title of Sade's hit song—a smooth operator. Her expected utility function and indifference curves are smooth, and when it comes to navigating the challenging choice situations that often confound *Homo sapiens*, she smoothly averts the pitfalls. Nevertheless, as we pivot to learning about the behavioral economic theories that have emerged to explain the divergences in choice behavior between the two species, bear in mind that to a great extent the reason why *Homo economicus* seems so smooth is because the world explained by the rational choice model is perforce restrictive of many of the characteristics that define the human experience—emotion, sensation, distraction, a world whose complications seem to demand imperfection and roughness (as opposed to smoothness). As a result, *Homo economicus* couldn't help but perform well within a bubble of surreality. Fortunately, as we are about to learn, the adjustments to the rational-choice framework devised by behavioral economists are quite efficacious, not to mention eloquent.

STUDY QUESTIONS

Note: Questions marked with a "‡" are adopted from Cartwright (2014).

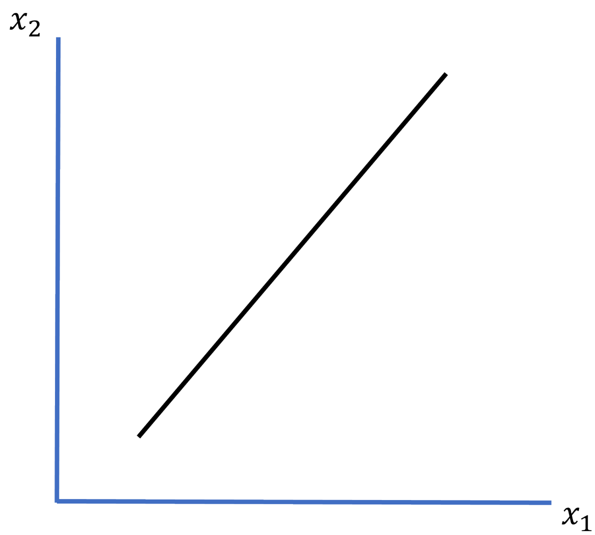
1. State the Completeness and Independence Axioms in two separate, easy-to-understand sentences.

2. Recall that Figures 3.1 – 3.3 are depicted for utility function $u\left(w_i\right) = \sqrt{w_i}$, indicating that *Homo economicus* exhibits "risk aversion" with respect to total wealth level . (a) Suppose instead that $u\left(w_i\right) = w_i$. Reconstruct Figures 3.1 – 3.3 for this utility function and interpret your results in relation to the figures in the text derived for $u\left(w_i\right) = \sqrt{w_i}$. Can you think of any reason why $u\left(w_i\right) = w_i$ might not be representative of *Homo economicus*' preferences? Explain. (b) Now let $u\left(w_i\right) = w_i^2$. Reconstruct Figures 3.1 – 3.3 for this utility function and interpret your results in relation to the figures in the text derived for $u\left(w_i\right) = \sqrt{w_i}$ as well as those you derived in part (a) for $u\left(w_i\right) = w_i$. Can you think of any reason why $u\left(w_i\right) = w_i^2$ might not be representative of *Homo economicus*? Explain.

3. ‡ Suppose Henry the *Homo economicus* is presented with the following lotteries: Lottery **A**: Win $6,000 with a probability of 0.45. Lottery **B**: Win $3,000 with a probability of 0.9. Lottery **C**: Win $6,000 with a probability of 0.001. Lottery **D**: Win $3,000 with a probability of 0.002. Show why it would be inconsistent for Henry to choose Lottery **B** over Lottery **A** together with Lottery **C** over Lottery **D**. What pattern do you notice in the four

lotteries?

4. In Figures 3.4 – 3.5, *Homo economicus*' indifference curve is drawn downward-sloping. Suppose instead that the indifference curve is upward-sloping, as depicted in the figure below. Interpret this figure. Hint: One of the two goods is actually a "bad" rather than a "good." Would *Homo economicus* ever exhibit an upward-sloping indifference curve?

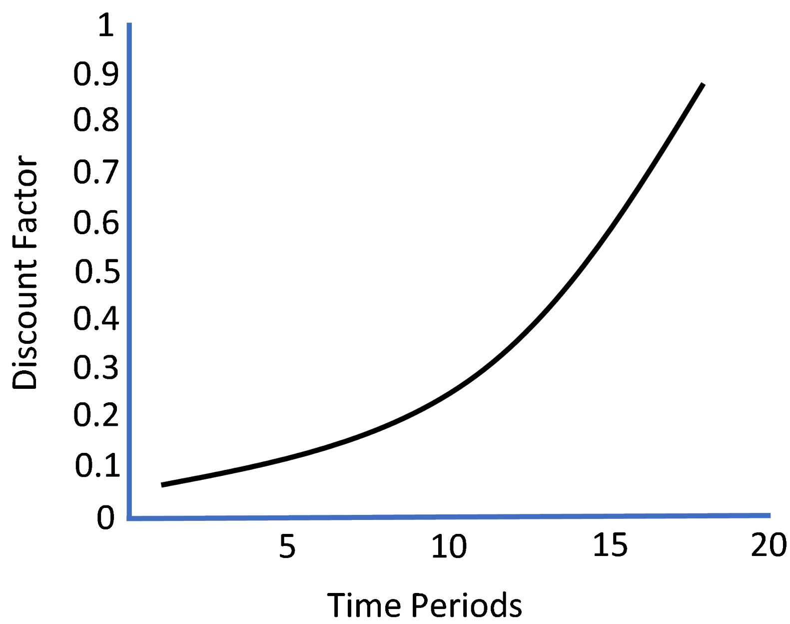


5. Given the upward-sloping indifference curve in Question 4, does satisfying the Transitivity Axiom still imply that indifference curves cannot intersect one another? Explain.

6. What does satisfying the Invariance Axiom imply about *Homo economicus*' susceptibility to Framing Effects studied in Chapter 1?

7. What does satisfying the Sure-Thing Principle imply about *Homo economicus*' susceptibility to Priming Effects studied in Chapter 1?

8. What would you call an individual whose time discounting path is depicted in the figure below (i.e., it fits the shape of a positive exponential curve rather than negative exponential curve as depicted in Figure 3.6)?



Media Attributions

- Figure 1 (Chapter 3) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 8 (Chapter 2) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 9 (Chapter 2) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 10 (Chapter 2) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 3.1 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 3.2 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 3.3v3
- Figure 3.4 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 3.5 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 3.6 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 4 (Chapter 3) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 8 (Chapter 3) © Arthur Caplan is licensed under a CC BY (Attribution) license

CHAPTER 4.

THE REALITY OF *HOMO SAPIENS*

You have already tested the reality of being a member of *Homo sapiens* in Chapters 1 and 2 by engaging in a variety of thought experiments and learning second-hand about experiments that have measured the extent to which people like you and me fall victim to effects such as Depletion, Priming, and Conformity, to name a few. Now it is time for you to engage in the same laboratory experiments that Kahneman, Tversky, Thaler, and others famously devised so that you can test just how far *Homo sapiens* deviate from the rationality axioms and other thresholds of consistency in our choice behavior. Before diving into the experiments though, we need to discuss (at some length) Kahneman and Tversky's (1979) revision to the expected utility theory presented in Chapter 3, which they called Prospect Theory. This is behavioral economics' bedrock theory. Making this detour here will enable us to set some crucial benchmarks for the experiments to follow.

PROSPECT THEORY**

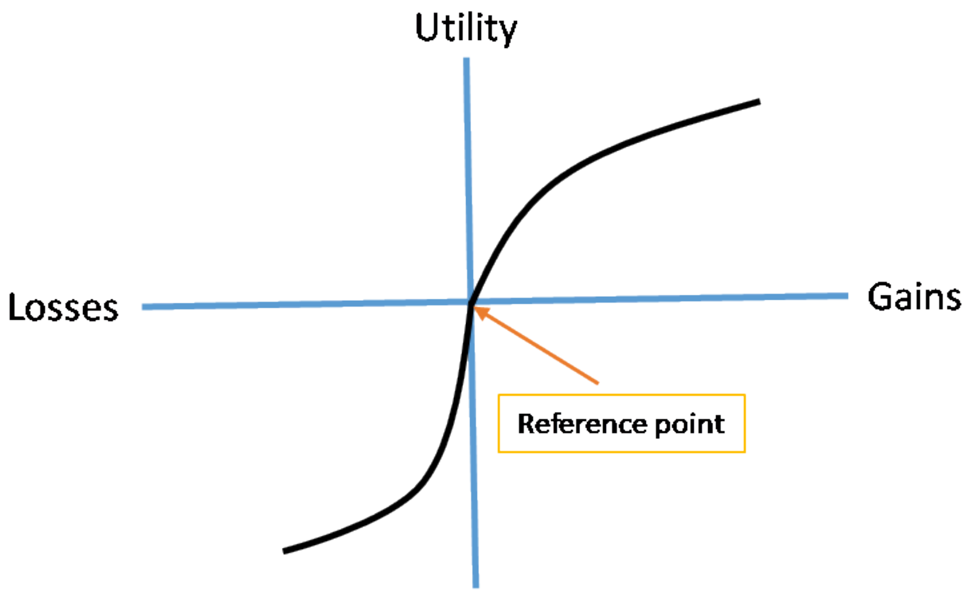
Several of the departures from the traditional rational choice model featured in Kahneman and Tversky's (1979) Prospect Theory conveniently arise in what appear to be innocuous adjustments to our original graph of utility function $u\left(w_i\right)$ depicted in Figure 3.1. As we will see, these adjustments are nuanced, so be careful not to jump to conclusions.[1]

For example, Kahneman and Tversky (1979) propose that people do not normally consider relatively small outcomes—like the wins and losses of the lotteries we've previously encountered—in terms of their total wealth, but rather in terms of the lottery's gains and losses independent from their initial wealth level. And just as an individual's utility can be represented as a concave function of the size of a gain from a lottery, the same can be said of a loss (i.e., the difference in (dis)utility between a loss of $200 and a loss of $100 appears greater than the difference in (dis)utility between a loss of $1,200 and a loss of $1,100). And to the extent that people suffer from "loss aversion," the concave function defined over losses is steeper than that defined over gains (i.e., *Homo sapiens* consider a loss of $X more averse than an equal but opposite gain of $X is deemed attractive).

These adjustments to the standard utility function first depicted in Figure 3.1 are pictured below in Figure 4.1, resulting in the individual's "value function."

Figure 4.1. *Homo sapiens*' Value Function (Prospect Theory)

1. Remember our discussion about Jumping to Conclusions earlier in the book?



Begin by noticing that the "reference point" for the value function is not the individual's initial wealth level.[2] Rather it is the origin of the graph, here corresponding to \$0. Next, as mentioned above, note that utility derived from gains, or a lottery's winnings, is concave just as it is for our original utility function $u(w_i)$. Thus, the value function similarly depicts diminishing sensitivity to gains. Finally, note that the individual's disutility derived from a lottery's losses is not only concave (thus depicting diminishing sensitivity to losses), but is also everywhere steeper, reflecting loss aversion.[3] As you will see later in Section 4, reference dependence and loss aversion have played heavily in subsequent empirical research and choice infrastructure in the field of behavioral economics.[4]

One immediate implication of the value function's shape in Figure 4.1 is that it is always best for *Homo sapiens* to aggregate losses but segregate gains. As an example, suppose *Homo sapiens* Sally suffers two distinct losses in a single evening out on the town but also enjoys two distinct gains. The losses are (1) the \$50 concert ticket she lost somewhere on the way to the theater, which she only realized was lost at the theater's entrance; and (2) the extra \$20 she had to pay to park closer to the theater that evening because she was running late. Her two gains from the evening were (1) the \$20 bill she happened to find on the sidewalk out in front of the theater; and (2) the \$25 bill at the restaurant that her friends paid for because they felt bad about Sally's losses from earlier that evening.

According to the value function, Sally will experience less disutility from her losses if she thinks of them in the aggregate—as a single \$70 loss—rather than two separate losses of \$50 and \$20. Because the value function is concave shaped in the loss region, the disutility corresponding to a \$70 loss is less than the total disutility corresponding to separate losses of \$50 and \$20. Using similar reasoning,

2. Reference point is sometimes referred to as an anchor or saliency point.

3. The ratio of the slopes of the loss and gain portions of the value function measured near the function's origin is a formal measure of loss aversion. Empirical estimates of loss aversion are typically close to a value of two, meaning that the disutility associated with an incremental loss is twice as great as the utility associated with an incremental gain of the same magnitude (Tversky and Kahneman, 1991, and Kahneman et al., 1990).

4. To whet your appetite, Odean (1998) used transaction data for 10,000 customers of a discount brokerage firm to examine how loss aversion might manifest itself in the trading decisions made by small-time investors in the stock market. He found that investors were much more likely to sell investments that had increased rather than decreased in value. This tendency to realize gains and avoid realizing losses (i.e., being averse to loss) is known as the Disposition Effect.

Sally will experience more utility from her gains if she thinks of them separately—as separate gains of \$20 and \$25—rather than as a single gain of \$45. This is because of the value function's concave shape defined over utility obtained from gains.

Thaler (1985) tested this implication of the value function and found that whether *Homo sapiens* choose to integrate losses and segregate gains depends upon how the losses and gains are framed to them in an experiment. This phenomenon is known as "hedonic framing." Notwithstanding this experimental evidence, the implication's lesson is clear: *Homo sapiens* would do well to aggregate multiple losses into a single loss and to keep multiple gains separated. As we will see below, this form of mind control is an example of what is known as "mental accounting." Although behavioral economists generally believe that mental accounting leads to sub-optimal decision-making via distorting an individual's allocation of wealth across the consumption of goods and services (c.f., Just, 2013), the value function implies that there are indeed situations where *Homo sapiens* can use mental accounting to their advantage.[5]

Another closely related implication of the value function is what's known as "hedonic editing." This is a case where individuals who lose something that was recently gifted to them are better off considering the loss as an elimination of a gain (originally attained from the gift) rather than as an outright loss. This is due to the steeper slope of the value function defined over losses relative to its slope over gains. Thaler and Johnson (1990) tested this implication in laboratory experiments over different periods of time and found no clear, convincing evidence that *Homo sapiens* engage in hedonic editing.

As a point of reference (not a reference point *per se*), Cartwright (2014) proffers a simple functional form for Kahneman and Tversky's (1979) value function depicted in Figure 4.1,

$$v(x;r,\alpha,\beta,\lambda) = \{ \begin{matrix} (x-r)^{\alpha} & if \quad x \geq r \\ -\lambda(r-x)^{\beta} & if \quad r > x \end{matrix}$$

where x represents the expected gain or loss experienced by an individual in relation to reference point r, and $\alpha > 0$, $\beta > 0$, and $\lambda > 0$ represent the function's additional parameters.[6] When the individual experiences a gain, $x \geq r$, the value function assigns a value of $(x-r)^{\alpha} > 0$ to that gain. And when the individual experiences a loss, the function assigns a value of $-\lambda(r-x)^{\beta} < 0$. To depict the value function in Figure 4.1, we must further refine the parameter values in $v(x;r,\alpha,\beta,\lambda)$. For those of you with strong enough math backgrounds, you will note that the depiction of Figure 4.1 requires $r = 0$, $0 < \alpha < 1$, $0 < \beta < 1$, and $\lambda > 1$. The additional restrictions on α and β (i.e., $\alpha < 1$ and $\beta < 1$) ensure diminishing sensitivity to gains and losses, and $\lambda > 1$ ensures loss aversion (λ is commonly known as the "coefficient of loss aversion"). These restrictions are revisited below when we briefly explore an alternative theory of *Homo sapiens* choice behavior known as Regret Theory.

Using results from a laboratory experiment with their students, Tversky and Kahneman (1992) were able to estimate parameters α, β, and λ; their median values being $\alpha = \beta$ = 0.88 and λ = 2.25. Further, the authors estimate decision weights that are calculated subjectively (and subconsciously) by *Homo sapiens* to adjust the objective probabilities associated with the gains and losses of any given lottery (note from Chapter 3 that objective probabilities p_g (for gain) and p_l (for loss) could be used

5. Interestingly, purchasing goods and services with a credit card and paying off the balance at the end of each billing period is a convenient way for individuals to aggregate losses (you pay for all the goods purchased that period in one, painful lump-sum) and segregate gains (you enjoy your purchases as you make them).

6. It is common notational practice to list a function's parameters to the right of the semi-colon and the variable over which the function is defined to the left of the semi-colon within the parenthesis.

by *Homo economicus* to calculate the expected value, x, of a given lottery as $x = p_g w_g + p_l w_l$, where w_g and w_l are the values of the lottery's gains and losses, respectively). Tversky and Kahneman (1992) estimated these decision weights, $\pi_i\left(p_i\right),\ i = g, l$, using the nasty-looking formula,

$$\pi_i(p_i) = \frac{p_i^{\gamma_i}}{\left(p_i^{\gamma_i} + (1-p_i)^{\gamma_i}\right)^{1/\gamma_i}}$$

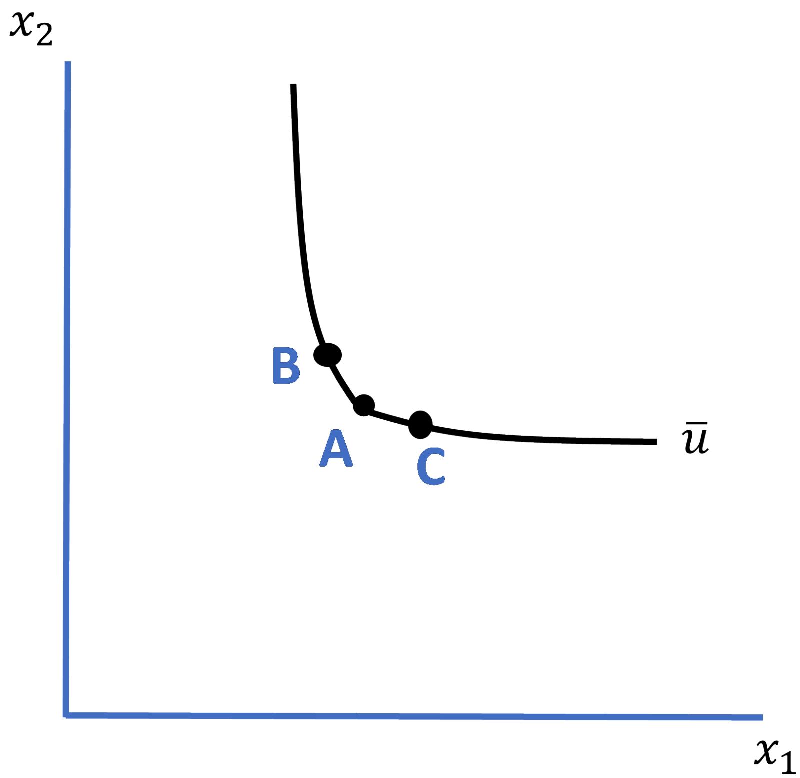
where the median value for parameters γ_i were estimated as γ_g = 0.61 and γ_l = 0.69. Thus, if the probabilities of gain and loss associated with a given lottery are p_g = 0.6 and p_l = 0.4, then using Tversky and Kahneman's median values for γ_g and γ_l results in decision weights of π_g = 0.47 and π_l = 0.39. Thus, in relation to their corresponding objective probabilities, *Homo sapiens*' subjective decision weights effectively reduce the lottery's probabilities associated with both the gains and losses, with the magnitude of the reduction applied to the probability of gain being larger than that applied to the probability of loss.

By way of example, if the lottery's gain is w_g = \$100 and loss is w_l = -\$105, then *Homo economicus* would calculate the lottery's expected value as (0.6 x \$100) – (0.4 x 105) = \$18, while *Homo sapiens* would calculate the value as (0.47 x \$100) – (0.39 x 105) = \$6. Although our example here suggests that Tversky and Kahneman's decision weights end up underweighting the objective probabilities of a lottery, we explain in Chapter 6 how the decision weights actually overweight improbable events (e.g., for a lottery with p_g = 0.05 rather than 0.6, as in our example here).[7]

Do reference dependence and loss aversion as embodied by value function $v\left(x; r, \alpha, \beta, \lambda\right)$ have implications for *Homo sapiens*' indifference curve? In a word, yes. As Just (2013) shows, reference dependence and loss aversion imply the existence of an almost-imperceptible kink in an individual's indifference curve at the reference point. To see this, consider Figure 4.2 below where bundle A represents the individual's reference point. By the monotonicity property, we know that any move in the northeast direction from bundle A to a bundle including more of both goods 1 and 2 puts the individual on a higher indifference curve representing higher utility. To the contrary, any move to the southwest from bundle A to a bundle including less of both goods places the individual on a lower indifference curve representing lower utility. But what of moves to the northwest or southeast of bundle A?

Figure 4.2. *Homo sapiens*' Kinked Indifference Curve

7. A homework problem asks you to calculate the decision weight π_g associated with p_g = 0.05 using the above decision-weight formula.



Begin by considering moves from bundle A northwest and southeast to new bundles B and C, respectively, along the indifference curve, where bundles B and C lie in what's known as a "local neighborhood" of bundle A (i.e., very close to bundle A). Note that the amount of good 1 in new bundle B must have decreased relative to its amount in bundle A, and the amount of good 2 must have increased. This case is vice-versa for bundle C, where the amount of good 1 has increased and the amount of good 2 has decreased.

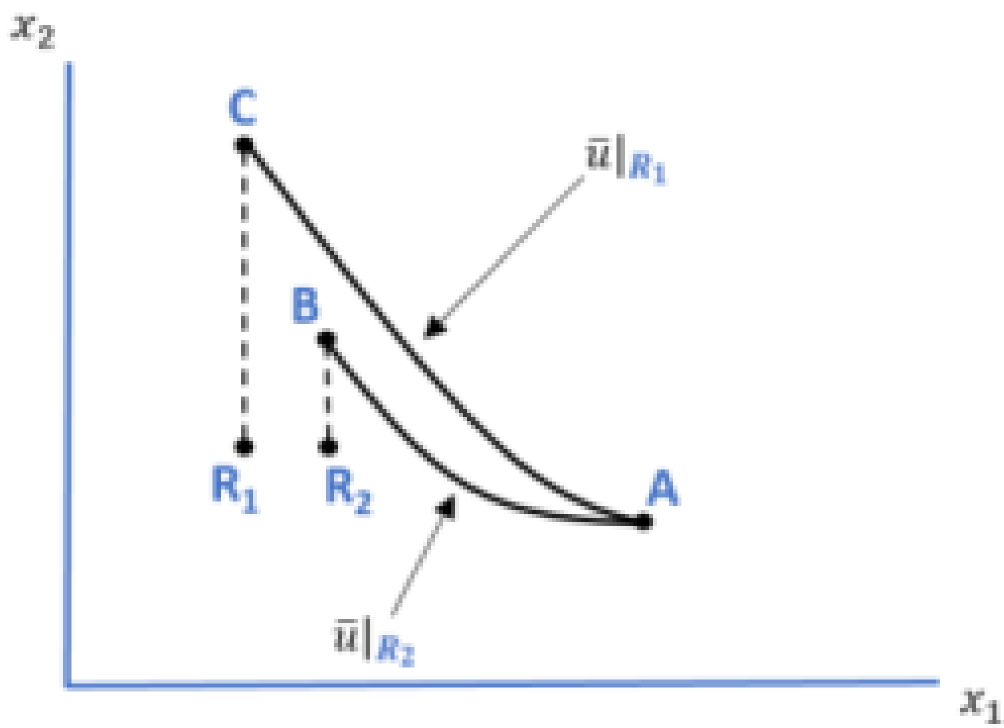
Now, because of *Homo sapiens*' proclivity for loss aversion, which you will recall manifests itself in a value function exhibiting an everywhere steeper curve defined over losses than over gains, it must be the case that the increase in bundle B's amount of good 2 is larger than its decrease in good 1. Otherwise, the individual would suffer a net loss in utility via loss aversion which, in turn, suggests that his utility would have to be represented by a lower (in the southwest direction) indifference curve rather than the one depicted by utility level $\bar{u}$ in the figure. By similar reasoning, it must be the case that bundle C's increase in good 1 is larger than its decrease in good 2. Otherwise, the individual would again suffer a net loss in utility via loss aversion. Thus, pulling these arguments together, the slope of the indifference curve between bundles A and B in Figure 4.2 must be larger than the slope of the curve between bundles A and C. While the stylistic, smoothly convex indifference curve typifies the preferences of *Homo economicus*, the stylistic indifference curve attributable to *Homo sapiens* exhibits tiny kinks throughout.

Furthermore, Just (2013) shows that if we reconsider the indifference curves for *Homo sapiens* as being reference-dependent themselves, then our family of reference-dependent indifference curves can accommodate intersections, unlike for *Homo economicus*, for whom reference dependence is a non-issue. Loss aversion is not necessary to obtain this result.[8]

8. As we will discuss, however, the family of reference-dependent indifference curves cannot be drawn to simultaneously abide by the Transitivity Axiom and Monotonicity Property. The family of curves continues to abide by monotonicity but not transitivity, and in this respect, suggests that *Homo sapiens* can indeed exhibit preference reversals (since preference

To see what happens when an individual's, say Jill's, indifference curves depict reference dependence, consider Figure 4.3 below where $\bar{u}|_{R_1}$ represents the utility level over which Jill's indifference curve is defined given reference bundle R1 and $\bar{u}|_{R_2}$ represents the utility level over which Jill's indifference curve is defined given reference bundle R2 (a mouthful, I know, but hopefully you can see this).[9]

Figure 4.3. *Homo sapiens*' Reference-Dependent Indifference Curves



Now consider bundle A in relation to reference bundles R1 and R2. Note that because reference bundles R1 and R2 each include the same amount of good 2, Jill gives up the same amount of good 2 by moving from either of the reference bundles to bundle A. In contrast, the amount of good 1 Jill obtains by moving from reference bundle R1 to bundle A is greater than the amount Jill gains of good 1 by moving from reference bundle R2 to bundle A. Do you see that?

And now comes the fun part—consider the following thought experiment. Starting with reference bundle R1, suppose Jill switches to bundle A and then is posed the following questions:

> "Jill, if you now had to sacrifice the increase in good 1 that you just obtained by switching from bundle R1 to bundle A, how much additional amount of good 2 would you require in order to retain utility level $\bar{u}|_{R_1}$? And if you instead started with reference bundle R2 and again

reversals are an implication of violating Transitivity). We wait until later in this section—after we have explored what's known as the Endowment Effect—to couch the implications of reference dependence, loss aversion, and the endowment effect directly into *Homo economicus*' family of non-reference-dependent indifference curves.

9. Because they are reference-dependent indifference curves rather than non-reference-dependent, the curves in this figure do not exhibit a kink as they did in the previous figure.

had to sacrifice the increase in good 1 that you just obtained by switching from bundle R2 to bundle A, how much additional amount of good 2 would you require in order to retain utility level $\bar{u}|_{R_2}$?"

Since the amount of good 1 gained in the move to bundle A is larger relative to bundle R1 than to bundle R2, the corresponding amount of good 2 Jill would require to maintain utility level $\bar{u}|_{R_1}$ likewise exceeds the amount of good 2 required for Jill to maintain utility level $\bar{u}|_{R_2}$ (via an application of the Monotonicity Property). Thus, connecting the dots, so to speak, Jill's reference-dependent indifference curve corresponding to reference bundle R1 (which includes points A and C) in Figure 4.3 is steeper than her curve corresponding to reference bundle R2 (which includes points A and B). *Voila*! The two indifference curves intersect at bundle A, which suggests a violation of the Transitivity Axiom from Chapter 3.

Again, because reference dependence is a non-issue for *Homo economicus* , reference-dependent indifference curves as drawn in Figure 4.3 are a non-starter for them. When it comes to indifference curves describing *Homo economicus*, we are restricted to drawing curves like those depicted in Figure 3.4.

REGRET THEORY**

As an alternative to Kahneman and Tversky's Prospect Theory, Loomes and Sugden (1982 and 1987) propose what they have called Regret Theory; a theory with implications for both the expected utility theory ascribed to *Homo economicus* and Prospect Theory's main constructs, in particular the causes of preference reversals (you will learn more about preference reversals in Chapter 5).

Regret Theory posits that an individual's preferences for a given lottery depends explicitly upon the other "unchosen" lottery (i.e., the tradeoff the individual makes between the chosen and unchosen lotteries). As a point of reference, recall *Homo economicus*' original utility function defined over wealth level w, $u(w)$, associated with a lottery ultimately chosen by the individual, say lottery L. Similar to Prospect Theory, which extended utility function $u(w)$ to value function $v\left(x; r, \alpha, \beta, \lambda\right)$, Regret Theory, in its most general form, extends *Homo economicus*' utility function to *Homo sapiens*' "Regret Theory Utility Function" $u\left(x, y\right)$, where now the individual's utility is simultaneously defined over levels x from a chosen lottery, say lottery L, and y from the unchosen lottery L' (Just, 2013). While function $u\left(x, y\right)$ is still increasing in x (as it is for *Homo economicus*), it is decreasing in y. For example, if the two lotteries are such that the individual ultimately receives \$5 from chosen lottery L (i.e., x = \$5) when unchosen lottery L' would have yielded \$1 (y = \$1), he is happier than if y would have instead yielded anything greater than \$1 (y > \$1). As long as $y < x$, the individual "rejoices" with $u\left(x, y\right) > 0$. If instead $y > x$, the individual experiences "regret" with $u\left(x, y\right) < 0$. At the threshold where y = x, $u\left(x, y\right)$ = 0. Regret Theory proposes that *Homo sapiens* maximize the expectation of $u\left(x, y\right)$.

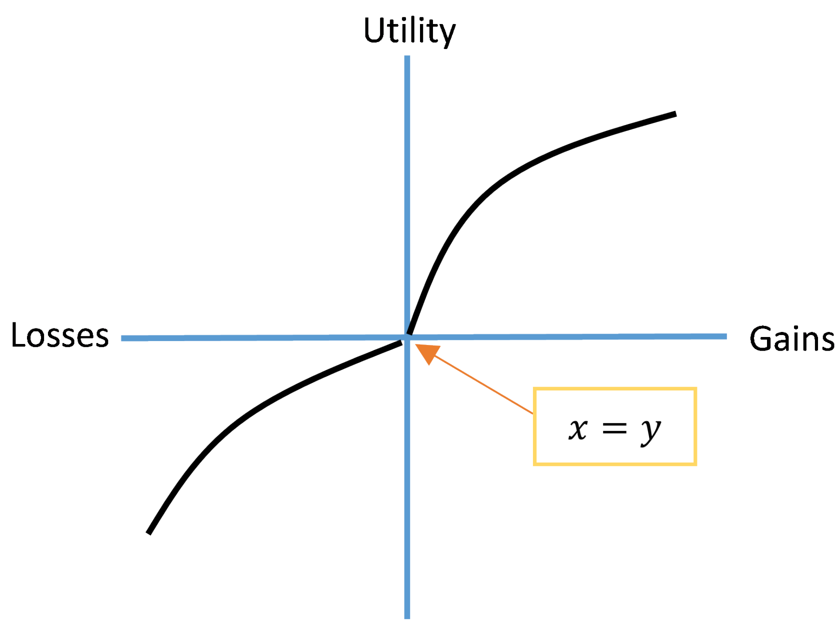
As Just (2013) explains, one version of the theory imposes symmetry on the individual's utility levels associated with rejoicing and regret (called skew symmetry) as well as super-additivity on the individual's utility function defined over regret (called regret aversion). Specifically, skew symmetry requires $u\left(x, y\right) = -u\left(y, x\right)$ along with $u\left(x, y\right)$ = 0 when x = y. Letting z represent the outcome of another unchosen lottery, regret aversion requires $u\left(y, x\right) < u\left(y, z\right) + u\left(z, x\right) < 0$ when the outcomes from the corresponding lotteries are $y > z > x$ (i.e., an individual regrets two small disappointments less than a single large one). With these properties in hand, Just (2013) goes on to show that Regret Theory can indeed justify preference reversals in the form of violations of the

Transitivity Axiom—that is, to the extent that it accurately depicts the emotion of regret suffered by *Homo sapiens*, Regret Theory admits preference reversals, which are anathema to the rational-choice dictates of *Homo economicus*.

It should come as no surprise that Regret Theory Utility Function admits preference reversals. Similar to what we showed earlier in Figure 4.3, reference-dependent indifference curves intersect. The Regret Theory Utility Function formalizes reference dependence via explicitly including the outcome associated with the unchosen lottery, y, as a variable in the function. Outcome y in turn acts as the individual's reference point.

Indeed, we can cast $u\left(x, y\right)$ directly in terms of value function $v\left(x; r, \alpha, \beta, \lambda\right)$ because of this correspondence between reference points r from Prospect Theory and y from Regret Theory—in effect $r = y$. Further, we can interpret skew symmetry as implying $\lambda = 1$ and regret aversion as implying $\beta > 1$ (as opposed to $\lambda = 1$ and $0 < \beta < 1$, respectively, for Prospect Theory according to Cartwright's (2014) simple functional form for the value function). As a result, the value function depicted in Figure 4.1 is recast for Regret Theory as Figure 4.4.

Figure 4.4. *Homo sapiens*' Value Function (Regret Theory)



Note two things about the Regret Theory version of the value function in this figure. First, the portion of the curve defined over losses is convex-shaped rather than concave-shaped, as it is for the Prospect Theory version of the curve (recall Figure 4.1). Second, the portion of the curve defined over losses is not necessarily more steeply sloped than the portion defined over gains. This is because Regret Theory does not propound the notion of loss aversion.

HOMO SAPIENS AND INTERTEMPORAL CHOICE***

In Chapter 3 we learned that when confronted with an intertemporal choice problem, *Homo economicus* maintains stable preferences over consumption goods, adopts exponential time discounting, and as a result, exhibits stationarity in choice comparisons over time.[10] As a result,

10. Recall that stationarity means that preferences for any increments of consumption in two different time periods depends upon the interval of time that passes between the two time periods (e.g., between periods 1 and 2), and not the specific points in time when the two respective increments could be consumed (e.g., periods 2 and 3).

Homo economicus makes time-consistent choices in terms of her consumption profile (and, by default, savings profile) over time. In other words, the consumption profile chosen by *Homo economicus* at the outset of her intertemporal decision problem does not change as she progresses from period to period and effectively re-solves her decision problem from each period forward. Clearly, from both theoretical and practical standpoints, any number of complications—embodied as relaxations of our model's underlying assumptions—could lead to an appearance of time-inconsistent choices. Foremost among these assumptions are perfect foresight and the constancy of preferences, prices, and income, which were described in Chapter 3. However, to level the proverbial playing field of comparison between the intertemporal choice problems of *Homo economicus* and *Homo sapiens*, we retain these assumptions. The only assumption we dispense with is exponential time discounting and with it the condition of stationarity.

Based upon a simple laboratory experiment, Thaler (1981) found that when offered one apple today or two tomorrow, most subjects choose one today—an extra day is too long to wait to receive only one additional apple. However, if asked the same question about one apple a year from now or two apples one year and one day from now, the subjects' preferences often reverse—they prefer to wait the extra day for the additional apple. After waiting a year, a day does not seem very long to wait to double consumption. Since the time interval between the choices is the same—one day in each case—this occurrence violates stationarity. *Homo economicus* would instead choose one apple today and one apple a year from now.

To capture this phenomenon—this violation of stationarity—among *Homo sapiens*, we replace the exponential time discounting practiced by *Homo economicus* with what's come to be known as "hyperbolic time discounting" (Ainslie, 1992).[11] Recall from Chapter 3 that exponential time discounting is based upon constant discount factor δ, resulting in the set of discount factors over time,

$$\delta(t) = \delta^t, \ t = 0, \ldots, T,$$

where time period $t = 0$ corresponds to the initial period and $t = T$ represents the final period, which in Chapter 3 was period 3, but which, in theory, could extend to $T = \infty$. Following Just (2014), hyperbolic discounting leads to a corresponding set of discount factors over time defined by,

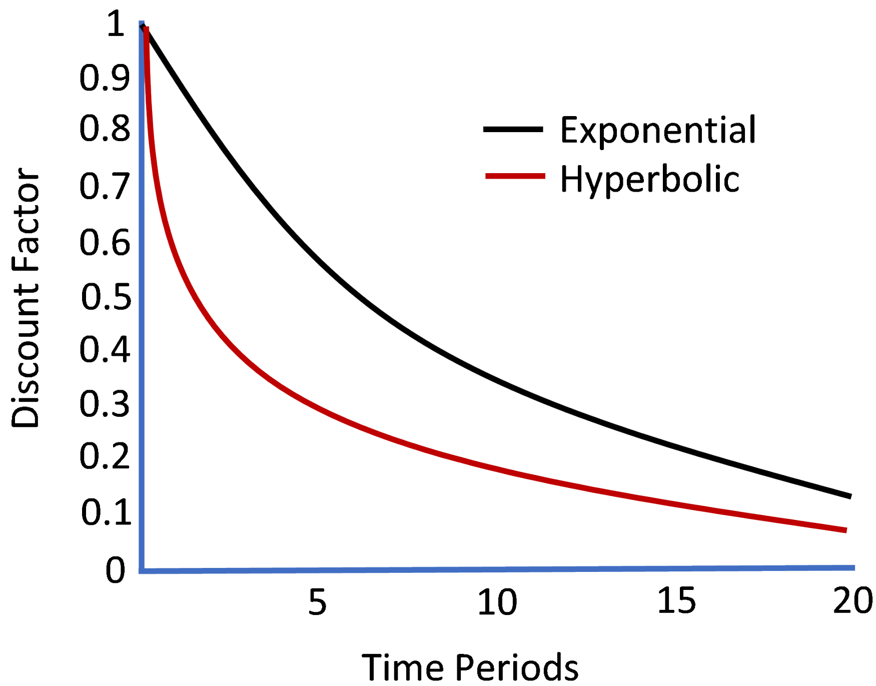
$$h(t) = (1 + \alpha t)^{-(\beta \alpha)}, \ t = 0, \ldots, T,$$

where scalars $\alpha > 0$ and $\beta > 0$. Both the exponential and hyperbolic discounting functions are depicted in Figure 4.5 for the case of $\alpha = \beta = 1$ and $\delta < 1$. Compared with the exponential discounting function, the hyperbolic discount factor declines more quickly over the first few time periods. Thus, an individual who abides by hyperbolic discounting values near-term future consumption much less than someone who abides by exponential discounting. The hyperbolic discounting function also declines very slowly over the latter periods relative to the exponential discounting function. Thus, an individual who abides by hyperbolic time discounting is more willing to delay consumption in the distant future than in the near future relative to an individual who discounts exponentially. In terms of Thaler's (1981) apple experiment, these discounting functions suggest that while subjects abiding by exponential and hyperbolic discounting might both prefer an

11. Laibson (1997) proposed a more tractable version of hyperbolic discounting that eventually became known as "quasi-hyperbolic discounting." As Cartwright (2014) explains, quasi-hyperbolic discounting can fully account for time-inconsistent behavior in *Homo sapiens*.

apple today rather than two apples tomorrow, subjects using hyperbolic discounting may end up preferring two apples a year and a day from now than one apple one year from now.[12], [13]

Figure 4.5. Hyperbolic vs. Exponential Time Discounting



Hence, *Homo sapiens*—who, as Thaler (1981) proposes, tend to adopt hyperbolic time discounting—do not display intertemporal preferences that adhere to stationarity. To see this formally, we borrow the notation from Chapter 3 for the individual's three-period choice decision and assume that $\alpha = \beta = 1$ in hyperbolic discounting function $h\,(t) = (1+\alpha t)^{-(\beta\alpha)}$, resulting in,

$$\underset{\{x_1, x_2, x_3\}}{Max} \quad u\,(x_1) + \tfrac{u(x_2)}{2} + \tfrac{u(x_3)}{3},$$

subject to,

$$W = (p_1 \cdot x_1) + (p_2 \cdot x_2) + (p_3 \cdot x_3),$$

where, for discounting purposes, we treat period 1 as equaling zero in the discounting function, and periods 2 and 3 as equaling one and two, respectively. Again, for those of you familiar with calculus, in particular solving constrained optimization problems, recall that you can write this problem in its Lagrangian form as,

$$\mathcal{L} = u\,(x_1) + \tfrac{u(x_2)}{2} + \tfrac{u(x_3)}{3}$$

where $\lambda > 0$ represents the problem's Lagrangian multiplier, and again for simplicity, we have normalized all prices to one (i.e., $p_1 = p_2 = p_3 = 1$). Obtaining the associated system of first-order conditions for this problem results in $u'\,(x_1) = u'\,(x_2)\,/2 = u'\,(x_3)\,/3 = \lambda$, where $u'\,(\cdot)$ denotes the individual's marginal utility function. Similar to *Homo economicus*' choice problem from Chapter 3,

12. Recall that if an individual abiding by exponential discounting prefers an apple today rather than two apples tomorrow, she will also prefer an apple a year from now more than two apples a year and a day from now.

13. See Benzion et al. (1989) for a classic laboratory experiment designed to infer individuals' time discounting regimes from choice scenarios involving (1) the postponement of a payment due (i.e., debt) until a later point in time, (2) postponement of a payment receipt (i.e., credit) until a later point in time, (3) expedition of a debt due in a future period to the current period, and (4) expedition of a credit expected in the future to the current period.

this string of equalities indicates that discounted marginal utility levels are equated across time. Given our underlying assumption of diminishing marginal utility, the string of equalities again implies $x_1^* > x_2^* > x_3^*$, i.e., the individual optimally consumes more in the first period than in the second, and more in the second period than the third.

To test *Homo sapiens* for stationarity, we follow the same approach used to test *Homo economicus* in Chapter 3:

> Suppose a member of *Homo sapiens* chooses to consume the same base amount x in each period and let two different increments to consumption be denoted as i_1 and i_2. Stationarity implies that if $u(x+i_1) > u(x+i_2)/2$, where $u(x+i_1)$ represents the corresponding utility level in period 1 and $u(x+i_2)/2$ represents discounted utility level in period 2, then $u(x+i_1)/2 > u(x+i_2)/3$ by exactly the same amount as $u(x+i_1) > u(x+i_2)/2$, where $u(x+i_2)/3$ represents discounted utility level in period 3. However, unlike what was shown for *Homo economicus* under exponential time discounting in Chapter 3, here we see that for *Homo sapiens* under hyperbolic time discounting, $(u(x+i_1)/2 - u(x+i_2)/3)$ does not equal $(u(x+i_1) - u(x+i_2)/2)$. Hence, *Homo sapiens*' preferences are not stationary through time.

Lastly, to see why *Homo sapiens*' preferences are potentially time inconsistent with hyperbolic discounting, we follow the same approach as was used to test *Homo economicus* for time consistency.[14] Given that he has chosen $x_1^*,\ x_2^*$, and x_3^* at the outset for any given set of δ, W, and prices p_1, p_2, and p_3, if after having consumed at level x_1^* in period 1, the individual decides to re-solve his decision problem from that point forward (i.e., now starting in period 2), he effectively solves,

$$\underset{\{x_2, x_3\}}{Max} \quad u(x_2) + \tfrac{u(x_3)}{2},$$

subject to,

$$W = 2w = (p_2 \cdot x_2) + (p_3 \cdot x_3),$$

which does not necessarily result in the same x_2^* and x_3^* as before. To see this result, first pull $u'(x_2)/2 = u'(x_3)/3$ from the string of three equalities derived from the individual's original decision problem (i.e., $u'(x_1) = u'(x_2)/2 = u'(x_3)/3$). Next, note that $u'(x_2)/2 = u'(x_3)/3$ is not the same equality resulting from the first-order conditions for this two-period problem, which is $u'(x_2) = u'(x_3)/2$. Therefore, given that nothing else has changed in this problem (i.e., the values for δ, w, and p_1, p_2, and p_3 are the same, as is the functional form of $u(\cdot)$), the values for x_2 and x_3 that solve *Homo sapiens*' two-period problem (here denoted as x_2^{**} and x_3^{**}, respectively) are not the same as the values for x_2^* and x_3^* that have solved his original three-period problem, specifically, $x_2^{**} > \ x_2^*$ and $x_3^{**} < \ x_3^*$. Thus, *Homo sapiens*' consumption profile is potentially time-inconsistent.

To see this last result, we first rewrite the equality from the individual's three-period problem, $u'(x_2)/2 = u'(x_3)/3$, as $u'(x_2) = 2u'(x_3)/3$ and then compare directly with the equality from the individual's two-period problem, $u'(x_2) = u'(x_3)/2$. Because the right-hand side of the former equality (i.e., $2u'(x_3)/3$) is larger than the right-hand side of the latter equality (i.e., $u'(x_3)/2$), if we take the x_2 and x_3 that solve the first equality (i.e., x_2^* and x_3^*) and plug these values into the second equality, then that equality no longer holds, in specific $u'(x_2^*) > u'(x_3^*)/2$. Given the assumption of diminishing marginal utility (i.e., that function $u'(\cdot)$ decreases in x), and that annual income levels

14. As Cartwright (2014) explains, quasi-hyperbolic discounting, not hyperbolic discounting, is necessary to (theoretically) account for time-inconsistent behavior among *Homo sapiens*. It is for this reason that we stress "potentially" time inconsistent in this sentence.

and all prices are fixed and constant, it therefore must be the case that $x_2^{**} > x_2^*$ and $x_3^{**} < x_3^*$ in order restore the equality.

You should note that this potential time-inconsistency result suggests that hyperbolic discounting provides a theoretical justification for why *Homo sapiens* are prone to procrastinate. If we think of the consumption good as embodying the value obtained from the feeling of relief that comes with putting forth the necessary effort to finish a task, then $x_2^{**} > x_2^*$ implies that the need to feel relief becomes more urgent (or valuable) by the time the individual reaches the second period. In other words, with hyperbolic discounting *Homo sapiens* undervalue the feeling of relief at the outset. As time goes by though (i.e., as he reaches the second period), *Homo sapiens* overcompensates his effort to obtain the feeling of relief. Sound like a recipe for procrastination? At the very least, hyperbolic time discounting is one of the recipe's main ingredients.

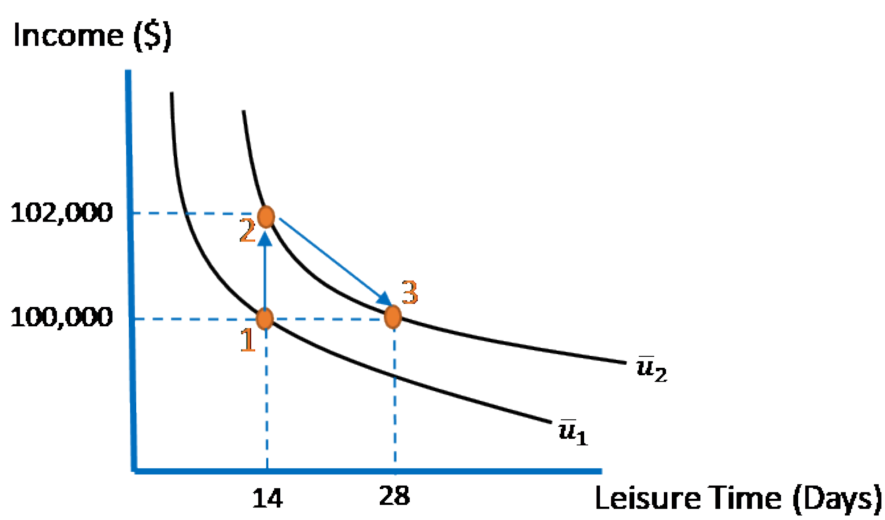
KEY TAKEAWAYS ON *HOMO SAPIENS*[**]

We conclude this chapter by revisiting the two central contributions of Kahneman and Tversky's (1979) Prospect Theory discussed earlier: reference dependence and loss aversion. We also introduce a new effect known as the Endowment Effect, which is a special case of the Anchoring Effect introduced in Chapter 1 and is also an expression of Status Quo Bias introduced in Chapter 2. As we will see in Chapter 5 and later in Section 4, the Endowment Effect has been the focus of several experiments.

Earlier, we relied upon a graphical representation of the value function to depict how the presence of reference dependence and loss aversion impelled a departure from the rational-choice model's conception of the expected utility form—the form used to represent *Homo economicus*' preferences over uncertain wealth and attendant risks. Here, following Tversky and Kahneman (1991), we depict all three idiosyncrasies in a single, indifference-curve framework.[15]

In Figure 4.6 below, the two "commodities" measured on the horizontal and vertical axes are, respectively, the number of vacation days and the annual income accruing to our exemplary individual, Tammy. Initially, Tammy is located at Bundle 1, earning $100,000 and taking 14 vacation days per year, and as a result, attaining $\bar{u}_1$ utils of happiness.

Figure 4.6. Reference Dependence, Loss Aversion, and the Endowment Effect



15. Kahneman et al. (1991) provide a nice synopsis of these key aspects of Prospect Theory.

Suppose Tammy is now given a raise (finally!) of \$2,000 per year. She moves to Bundle 2, attaining the higher $\bar{u}_2$ utils of happiness. A month later, out of the blue, her boss asks Tammy if she would be willing to forfeit the \$2,000 raise and instead take an extra 14 days of vacation time each year. If she accepts her boss' offer, she would therefore move to Bundle 3. Because she is on the same indifference curve with Bundle 3 as she is with Bundle 2, Tammy would still attain $\bar{u}_2$ utils of happiness.

To the extent that Tammy behaves more like a *Homo economicus* than *Homo sapiens*, what we would expect Tammy to do—hang on to her raise and forgo the extra 14 days of vacation time (i.e., remain at Bundle 2), or forfeit her raise and nab the extra vacation time (i.e., move to Bundle 3)? As a member of *Homo economicus*, Tammy would not suffer from reference dependence or loss aversion in moving from Bundle 2 to Bundle 3. And in this case, because her utility level remains constant at $\bar{u}_2$, neither would she exhibit an endowment effect. An Endowment Effect occurs when an individual perceives herself as better off in the status quo (e.g., at bundle 2) even when her utility level in the status quo is no higher than in a different state of the world (e.g., bundle 3). As a result, Tammy would be no worse off flipping a fair coin and letting the outcome of the coin flip determine her choice (e.g., "heads I stick with Bundle 2, tails I change to Bundle 3").

This would not be so if Tammy is a member of *Homo sapiens*. As a *Homo sapiens*, we would expect her to be partially governed by all three idiosyncrasies. Regarding reference dependence, consider the fact that Tammy has only enjoyed her raise in pay for one month. This bump in her income is therefore likely to be fresh in her mind. To the extent that this is the case, it could be that in the interim of having received the pay raise and being given the opportunity of choosing Bundle 3, her indifference curve associated with utility level $\bar{u}_2$ has actually gotten flatter in the region between Bundles 2 and 3 (with Bundle 2 serving as a pivot point), and thus, choosing Bundle 3 would now result in her attaining a utility level less than $\bar{u}_2$ (sketch this possibility in Figure 4.6 and see for yourself). Tammy therefore would not flip a coin. She would turn down her boss' offer and stick with Bundle 2.

Alternatively, it could be that Tammy's reference point for this decision is stuck at Bundle 1. One month hasn't really been long enough for her to learn to enjoy the added utility that she will eventually obtain from the pay raise. So, in her mind, Tammy actually compares Bundles 1 and 3, not Bundles 2 and 3. In this case, Tammy will accept the boss's offer and switch to Bundle 3. As far as Tammy in concerned, she has gained $\bar{u}_2 - \bar{u}_1$ utils of happiness in making the switch.

Either way, therefore, reference dependence nuances Tammy's decision when she thinks more like *Homo sapiens* than *Homo economicus*.

The story is less ambiguous regarding loss aversion. To the extent that Tammy suffers from loss aversion, she will interpret the certain loss of her \$2,000 pay raise as inducing a greater loss in utility than the potential gain in happiness that will come with 14 more vacation days. Hence, loss aversion points Tammy toward rejecting her boss' offer and sticking with Bundle 2 in Figure 4.6.

Finally, let's consider the possibility of Tammy experiencing an Endowment Effect as a member of *Homo economicus*. Because the move from Bundle 2 to Bundle 3 does not change her utility level, we would not expect Tammy, as *Homo economicus*, to suffer from an endowment effect (i.e., to necessarily choose Bundle 2 over Bundle 3). However, as a member of *Homo sapiens*, the Endowment Effect is potentially alive and well. Tammy could be covetous of her recent pay raise, and therefore would not be indifferent between Bundles 2 and 3. She would be more likely to stick with Bundle 2.

Tallying up the score in Figure 4.6 (and remembering that Tammy is, after all, a member of *Homo sapiens*), it seems that, contrary to what rational-choice theory would suggest, the three idiosyncrasies—reference dependence, loss aversion, and the Endowment Effect—point Tammy toward rejecting her boss' added-vacation offer and sticking with Bundle 2. At the very least, we should not be as confident in concluding that she will base her decision on a flip of a coin, as we would if she were miraculously a member of *Homo economicus* instead.

We conclude with a few quick thought experiments before taking a headlong dive into the famous laboratory experiments that have propelled the field of behavioral economics into the limelight.

Consider the following thought experiment:

> Today, Sally and Sam each have wealth worth $5 million. Yesterday, Sally had wealth worth $1 million and Sam had wealth worth $9 million. Are Sally and Sam equally happy today?

If you answered "no," then you believe that Sally and Sam's utilities are each reference dependent. By default, *Homo economicus* would answer "yes" because what matters in the rational choice model is total wealth (which is currently at the same level for Sally and Sam), not actual gains and losses.

Now consider this thought experiment:

> Which lottery do you prefer?
>
> **A** 85% chance to lose $1000 and 15% chance to lose nothing
> **B** certain loss of $800

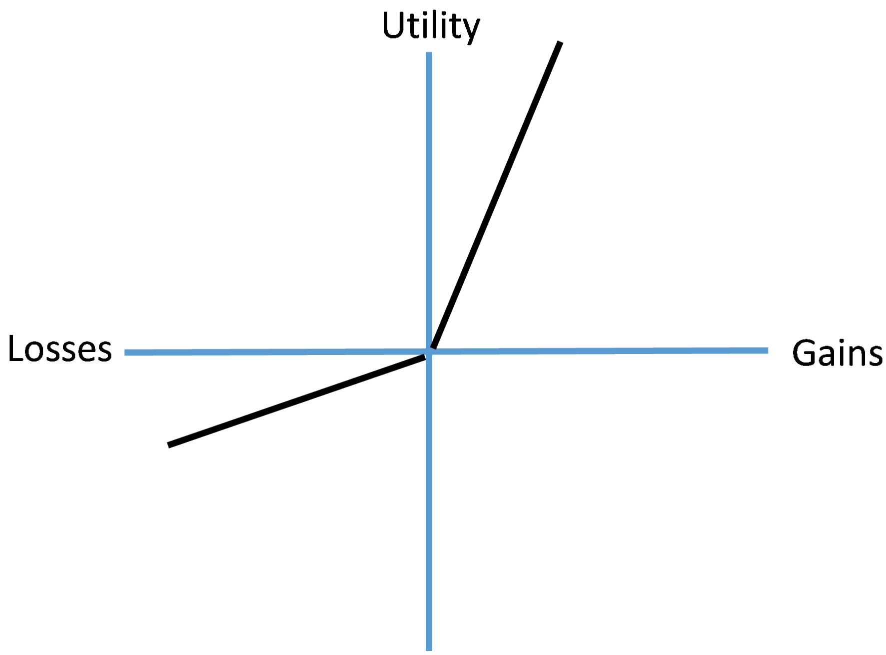
If you chose lottery A, then you are a "risk seeker" and exhibit loss aversion. *Homo economicus* would have done the math, and since her expected loss from lottery A is a larger number than the certain loss from lottery B, she would have chosen lottery B.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013), and those marked with a "‡" are adopted from Cartwright (2014).

1. † Suppose Uncle Joe's preferences can be depicted as a value function like that drawn in Figure 4.1. Last night Joe experienced both a gain and a loss. The gain was paying $25 less than he had expected for his dinner date with Auntie Jill. The loss was finding a $25 parking ticket waiting for him underneath his car's wiper blade when he and Jill returned to the car following the meal. Would it be best, in terms of his overall utility level, if Joe segregated the gain from the loss or integrated the two? Explain.

2. Suppose an individual's value function is depicted in the figure below. Which properties of Prospect Theory is this individual violating?



3. ‡ Suppose someone's willingness-to-pay for a good is \$10, and their reference point is \$20. If the good is priced at \$13, will they buy it? What does your answer say about sales and "bargain buys"?

4. Is the Monotonicity Property necessary for an individual's reference-dependent indifference curves to cross? Why?

5. † Suppose Akira has two sources of income. Anticipated income (e.g., accrued from her regular weekly paycheck), w_1, is spent on healthy food, x_1, and clothing, x_2. Unanticipated income (e.g., inherited wealth), w_2, is spent on what Akira considers to be a luxury good, x_3. Suppose Akira's value function is given by $v\left(x_1, x_2, x_3\right) = \left(x_1 x_2 x_3\right)^{\frac{1}{3}}$, so that the marginal utilities associated with goods 1, 2, and 3, respectively, are given by $MU_1 = \frac{1}{3} x_1^{-\frac{2}{3}} \left(x_2 x_3\right)^{\frac{1}{3}}$, $MU_2 = \frac{1}{3} x_2^{-\frac{2}{3}} \left(x_1 x_3\right)^{\frac{1}{3}}$, and $MU_3 = \frac{1}{3} x_3^{-\frac{2}{3}} \left(x_1 x_2\right)^{\frac{1}{3}}$. Suppose w_1 = \$8 and w_2 = \$2, and that corresponding per-unit prices for goods 1, 2, and 3 are given by $p_1 = p_2$ = \$1 and p_3 = \$2, respectively. It can be shown that Akira's optimal demands for goods 1, 2, and 3 are values 4, 4, and 1, respectively. To see this, we set the marginal utilities of consumption per dollar equal across x_1 and x_2 and then impose the condition(s) that the cost of all goods in the budget associated with the anticipated and unanticipated incomes are equal to their respective budget constraints. Hence, from the budget constraints for anticipated and unanticipated income, respectively, we set $x_1 + x_2$ = 8 and $2 = 2x_3 \Longrightarrow x_3 = 1$. Setting $\frac{MU_1}{p_1} = \frac{MU_2}{p_2} \Longrightarrow x_1 = x_2$, which, given $x_1 + x_2$ = 8, implies $x_1 = x_2 = 4$. (a) Suppose Akira receives an extra \$4 in anticipated income, and thus w_1 = \$12 and w_2 = \$2. How does Akira's demand for goods 1, 2, and 3 change? (b) Alternatively, suppose Akira receives the extra \$4 as unanticipated income, and thus w_1 = \$8 and w_2 = \$6. How does Akira's demand

change now? Hint for parts (a) and (b): Use the same approach as was shown above to answer these two questions. (c) In relation to parts (a) and (b), could Akira make herself better off by not doing mental accounting (i.e., by not compartmentalizing the entire $4 increase into either anticipated or unanticipated income) and instead combine the two accounts into a single total income account and allocate the $4 increase in income to total income? Hint for part (c): Note that total income becomes $14, and the corresponding budget constraint is $14 = x_1 + x_2 + 2x_3$. Again, noting that $\frac{MU_1}{p_1} = \frac{MU_2}{p_2}$, and now $\frac{MU_1}{p_1} = \frac{MU_3}{p_3}$ as well, leads to the answer.

6. ‡ Suppose Anna owns her own house and house prices in the market increase. Should Anna spend more money on everyday living expenses? Suppose Anna also has money invested in the stock market and stock prices decrease. Should Anna spend less money on everyday living expenses? Explain.

7. † Consider the Regret Theory Utility Function given by $u(x,y) = \{ \begin{matrix} (x-y)^2 & if & x \geq y \\ -(y-x)^2 & if & y > x \end{matrix}$. Plot the implied indifference curves for $u(x,y)$ = 1 and $u(x,y)$ = -1. Explain how these curves can be used to denote a preference reversal.

8. Explain why reference dependence in the context of Regret Theory can lead to the intersection of the reference-dependent indifference curves and thus a preference reversal.

9. Referring to the three-period hyperbolic time discounting problem discussed in this chapter, show that stationarity does not hold for utility function $u\left(x\right) = \ln x$. Hint: Assume α = β = 1. To show this result, you must then choose increments i_1 and i_2 such that $\ln\left(x + \ i_1\right) > \frac{ln(x+\ i_2)}{2}$, and proceed from there.

10. ‡ Referring to the lotteries listed below, explain what Prospect Theory suggests Henrietta the *Homo sapiens* will choose to do vis-à-vis Lottery A vs. Lotteries B – F, respectively. Lottery **A**: Win $0 for certain. Lottery **B**: Lose $100 with probability of 0.5, win $105 otherwise. Lottery **C**: Lose $100 with probability of 0.5, win $125 otherwise. Lottery **D**: Lose $100 with probability of 0.5, win $200 otherwise. Lottery **E**: Lose $225 with probability 0.5, win $375 otherwise. Lottery **F**: Lose $600 with probability 0.5, win $36 million otherwise.

11. † Harper is spending a three-day weekend at her beach property. Upon arrival, she purchases a quart of ice cream and must divide consumption of the quart over each of the three days. Her instantaneous utility of ice cream consumption is given by $u\left(x\right) = \sqrt{x}$, where x is measured in quarts, so that the instantaneous marginal utility is given by $u'\left(x\right) = \frac{1}{2\sqrt{x}}$. (a) Suppose Harper discounts future consumption of ice cream using exponential time discounting with a daily discount factor δ = 0.8. (a) Solve for Harper's optimal consumption profile over the course of the three days by finding the daily proportions of the quart of ice cream that both equate the discounted marginal utilities of consumption across the three days, and sum to 1.

(b) Now suppose that Harper discounts future utility according to hyperbolic time discounting, with $\alpha = \beta = 1$. Describe the optimal consumption profile starting from the first day of the weekend. How will the consumption plan change on day two?

12. ‡ Suppose Henrietta the *Homo sapiens* is prone to experience regret when it comes to choosing between different lotteries. She faces the following three lotteries: Lottery **A**: Win \$5 for certain. Lottery **B**: Win \$10 with probability of 0.4, win \$3 with probability of 0.6. Lottery **C**: Win \$7.50 with probability of 0.7, win \$1 with probability of 0.3. Given what you know about Regret Theory, how might you explain Henrietta's choices when comparing Lotteries A and B, B and C, and A and C, respectively?

13. What condition typically exhibited by *Homo sapiens* do businesses attempt to exploit when they advertise that "supplies won't last long"?

14. ‡ Is it better to be an employee in a firm where you earn \$50,000 per year and the average salary is \$80,000 or in a firm where you earn \$45,000 and the average salary is \$30,000.

15. Why don't the reference-dependent indifference curves $\bar{u}|_{R_1}$ and $\bar{u}|_{R_2}$ in Figure 4.3 extend to the southeast of Bundle A?

Media Attributions

CHAPTER 5.

LABORATORY EXPERIMENTS: THE RATIONALITY OF *HOMO ECONOMICUS* VERSUS THE REALITY OF *HOMO SAPIENS*

The laboratory experiments discussed in this chapter have been designed to test the Principle and Additional Rationality Axioms presented in Chapter 3. As you will see, we *Homo sapiens* are rather prolific in our violations of the rationality typified by *Homo economicus.*

TESTING THE INVARIANCE AXIOM (VERSION 1)

Consider the following experiments designed by Kahneman and Tversky (1984):

Experiment 1

Imagine that your hometown is preparing for the outbreak of an unusual disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

If Program **A** is adopted, 200 people will be saved.

If Program **B** is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

Which of the two programs would you favor?

Experiment 2

Imagine that your hometown is preparing for the outbreak of an unusual disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

If Program **C** is adopted, 400 people will die.

If Program **D** is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die.

Which of the two programs would you favor?

Homo economicus would quickly determine that Program A in Experiment 1 is equivalent to Program C in Experiment 2, and Program B in Experiment 1 is similarly equivalent to Program D in

Experiment 2. Thus, he would recognize that the two experiments themselves are equivalent. If each experiment were conducted separately with two different groups of *Homo economicus*, we would then naturally predict 50%-50% splits between Programs A and B among subjects participating in Experiment 1, and 50%-50% splits between Programs C and D among subjects participating in Experiment 2. These outcomes would be consistent with the Invariance Axiom.

To the contrary, when Kahneman and Tversky ran the experiments, 72%(28%) of the subjects in Experiment 1 chose Program A(B). The reverse occurred in Experiment 2, with 22%(78%) of the subjects choosing Program C(D). The authors surmised that this violation of the Invariance Axiom among *Homo sapiens* resulted from the experiments having been framed by different reference points. Experiment 1's reference point is that people will be saved while Experiment 2's is that people will die. This, in turn, led the *Homo sapiens* to make reference-dependent choices. Kahneman and Tversky identified this particular type of reference-dependency as a Reflection Effect.

TESTING THE INVARIANCE AXIOM (VERSION 2)

Kahneman and Tversky (1979) provide another test of reference-dependency and the Reflection Effect in an experiment with two groups of roughly 70 subjects each. The first group participated in the following experiment:

Experiment 1

Suppose you've been given $1,000 in addition to whatever you own in your life. Which lottery do you prefer?

A 50% chance to win another $1,000 (with a 50% chance to win nothing).
B Certain win of $500.

Since the two lotteries have the same expected values of $500, a *Homo economicus* with any degree of risk aversion would choose Lottery B. Of Kahneman and Tversky's 70 *Homo sapiens* who participated in this experiment, 84% chose Lottery B. Not bad.

The second group of subjects participated in a slightly altered version of Experiment 1:

Experiment 2

Suppose you've been given $2,000 in addition to whatever you own in your life. Which lottery do you prefer?

A 50% chance of losing $1,000 (with a 50% chance of losing nothing).
B Certain loss of $500.

Note that these two lotteries are essentially identical to the two lotteries in Experiment 1. Lotteries A both give the individual a 50% chance of walking away with $2,000 and a 50% chance of walking away with $1,500. Lotteries B both ensure that the individual walks away $1,500 wealthier with certainty. Thus, any *Homo economicus* who would choose Lottery B in Experiment 1 should likewise choose Lottery B in Experiment 2 (or to put it another way, the percentages of *Homo economicus* choosing Lottery B in Experiments 1 and 2 should be equal). We would like to think that *Homo sapiens* will behave similarly.

Wouldn't you know it? Of Kahneman and Tversky's 70 *Homo sapiens* who participated in this experiment, only 31% chose Lottery B. What the…? This outcome led Kahneman and Tversky to conclude that their subjects were indeed making reference-dependent choices. Experiment 1 was framed in terms of winning, and the great majority of *Homo sapiens* responded by exhibiting risk aversion—they prefer to protect certain gains. In contrast, Experiment 2 was framed in terms of losing, and the majority of *Homo sapiens* responded by exhibiting risk-seeking behavior—they decided to take a gamble that was otherwise eschewed in Experiment 1. The results for Experiment 1 concur with the value function's diminishing sensitivity to gains as described in Chapter 4, and the results for Experiment 2 concur with the value function's diminishing sensitivity to losses.

TESTING THE INVARIANCE AXIOM (VERSION 3)

Consider the following two experiments proposed by Grether and Plott (1979):

Experiment 1

Suppose you are offered a choice between two lotteries. Which lottery do you prefer?

A If you roll a 1 or a 2 you win $160; if you roll a 3, 4, 5, or 6 you lose $15.
B If you roll a 1, 2, 3, 4, or 5 you win $40; if you roll a 6 you lose $10.

Experiment 2

Suppose you "own" each of the two lotteries below, and therefore you have the option of selling each of them to someone else rather than playing them yourself. How much money would you sell each one for?

A If you roll a 1 or a 2 you win $160; if you roll a 3, 4, 5, or 6 you lose $15.
B If you roll a 1, 2, 3, 4, or 5 you win $40; if you roll a 6 you lose $10.

In this case, each subject participates in both experiments. We would expect that if *Homo economicus* prefers Lottery A to Lottery B in Experiment 1, then she would choose to sell Lottery A for more money than Lottery B in Experiment 2. This would be consistent with the Invariance Axiom in the context of this experiment. By contrast, Grether and Plott found that 70% of the *Homo sapiens* who participated in the two experiments exhibited a preference reversal by choosing the relatively safe Lottery B in Experiment 1 but stating a higher selling price for risky Lottery A in Experiment 2.[1]

A less-complicated version of Experiments 1 and 2 designed to test for preference reversal in the context of a monetary bet was tested by Tversky et al. (1990), who proposed an experiment to their students similar to the following:

> Suppose you are asked to choose between the following two lotteries. Which lottery do you prefer?
>
> **A** 75% chance of winning $10.
>
> **B** 10% chance of winning $100.

The students were asked two questions: which lottery would you prefer to play, and which lottery is worth more to you (in terms of the minimum amounts of money you would be willing to accept in lieu of having the chance to play either lottery)?

We know how *Homo economicus* would answer. He would first calculate the expected value of each lottery ($7.50 for Lottery A and $10 for Lottery B) and then answer the two questions as "I prefer Lottery B" and "Lottery B is worth $2.50 more to me." In other words, *Homo economics* would go Lottery B all the way; no preference reversal there. To the contrary, Tversky et al. (1990) found that around 75% of their subjects chose Lottery A in answer to the first question but roughly 65% of these same subjects chose Lottery B in answer to the second question. Ouch, big-time preference reversal there.

TESTING THE INVARIANCE AXIOM (VERSION 4)

Consider the following three experiments, versions of which were proposed by Kahneman (2011):

> **Experiment 1**
>
> The Andaman Islands in the Bay of Bengal off the coast of Myanmar are home to several varieties of animals. The breeding grounds for one animal in particular, the Narcondam Hornbill, are threatened by human settlement and consequent deforestation. Suppose a special fund supported by private donations has been set up to provide protected breeding locations for the Hornbill. Would you consider contributing something to this fund? If so, how much?

1. Lichtenstein and Slovic (1971) found similar preference reversals in their earlier study with undergraduate students, as did Loomes et al. (1991) in later experiments.

Experiment 2

Farmworkers, who are exposed to the sun for many hours per day, have a higher risk of skin cancer due to climate change than Myanmar's general population. Frequent medical check-ups can reduce the risk. Suppose a special fund supported by private donations has been set up to provide regular medical check-ups for the farm workers. Would you consider contributing something to this fund? If so, how much?

Experiment 3

The Andaman Islands in the Bay of Bengal off the coast of Myanmar are home to several varieties of animals. The breeding grounds for one animal in particular, the Narcondam Hornbill, are threatened by human settlement and consequent deforestation. Farmworkers, who are exposed to the sun for many hours per day, have a higher risk of skin cancer due to climate change than Myanmar's general population. Frequent medical check-ups can reduce the risk. Suppose separate special funds supported by private donations have been set up to provide protected breeding locations for the Hornbill and regular medical check-ups for the farmworkers. Would you consider contributing to one or the other fund (or both)? If so, how much?

Suppose that subjects recruited to participate in these experiments are divided into two groups. One group participates in Experiments 1 and 2 simultaneously, i.e., the subjects are asked to contribute to a special fund for the Narcondam Hornbills in Experiment 1 and the farmworkers in Experiment 2. The other group participates in Experiment 3. What would we expect from two groups of *Homo economicus* here? If you answer that the average amounts *Homo economicus* subjects pledged in Experiments 1 and 2 equal the same average amounts in Experiment 3, then you've nailed it! Essentially, both groups are presented with the same experiments. Thus, on average, *Homo economicus* would violate the Invariance Axiom if the amounts pledged for Experiments 1 and 2 did not match those pledged in Experiment 3.

Kahneman hypothesizes that the average amount pledged by *Homo sapiens* in Experiment 1 (for Narcondam Hornbills) will exceed the average amount pledged in Experiment 2 (for the farmworkers). But in Experiment 3, the average amounts will be reversed, indicating a preference reversal.[2]

TESTING THE INVARIANCE AND DOMINANCE AXIOMS (VERSION 1)

Consider the following experiments proposed by Kahneman and Tversky's (1984):[3]

2. Kahneman explains that presenting subjects with two separate questions (i.e., Experiments 1 and 2) frames their choices narrowly. Presenting the subjects with a single question (i.e., Experiment 3) instead frames their choices broadly. In this case, as with most cases, the broader the frame the more likely subjects will provide accurate answers—accurate in terms of pledging amounts that more accurately reflect their underlying preferences.
3. Taken together, these two experiments exemplify the famous Allais Paradox designed by Maurice Allais in 1953.

Experiment 1

Choose between lotteries **A** and **B**:

A 25% chance to win $240 and 75% chance to lose $760, or
B 25% chance to win $250 and 75% chance to lose $750.

Experiment 2

Now suppose you face the following pair of what are known as “compound lotteries”:

Compound Lottery 1: Choose between,

A a sure gain of $240, or
B 25% chance to win $1,000 and 75% to win nothing.

Compound Lottery 2: Choose between,

A a sure loss of $750, or
B 75% chance to lose $1,000 and 25% chance to lose nothing.

Begin by noting that Lottery B in Experiment 1 dominates Lottery A. This is because the expected winnings from Lottery B are greater than those from Lottery A, and the expected losses from Lottery B are less than those from Lottery A. Obviously, *Homo economicus* will choose Lottery B, thus not violating the Dominance Axiom.

Next comes the hard part. In Experiment 2, adding the sure win of $240 (Lottery A in Compound Lottery 1) to Lottery B in Compound Lottery 2 yields a 25% chance of winning $240 and a 75% chance to lose $760. Note that this is exactly Lottery A in Experiment 1! Similarly, adding the sure loss of $750 (Lottery A in Compound Lottery 2) to Lottery B in Compound Lottery 1 yields a 25% chance to win $250 and a 75% chance to lose $750. But this is precisely Lottery B in Experiment 1!

Thus, since *Homo economicus* will choose Lottery B in Experiment 1, in Experiment 2 he will choose Lottery B in Compound Lottery 1 and Lottery A in Compound Lottery 2.[4]

What about *Homo sapiens*? Thankfully, none of Kahneman and Tversky’s subjects in Experiment 1 chose Lottery A, implying that *Homo sapiens* also abided by the Dominance Axiom. However, in Experiment 2, the great majority of subjects chose Lottery A in Compound Lottery 1 and Lottery B in Compound Lottery 2. This is the opposite of *Homo economicus*’ choices and demonstrates a preference reversal for these *Homo sapiens* (relative to their choices of Lottery B in Experiment 1). In other words, once again a majority of *Homo sapiens* have violated the Invariance Axiom.

Kahneman (2011) reminds us that Experiment 2 is an example of narrow vs. broad framing.[5] Narrow framing occurs when subjects consider the two compound lotteries separately from each

4. Note that by choosing Lottery A in Compound Lottery 2, Homo economicus demonstrates that he does not suffer from loss aversion.

5. Also known as narrow vs. broad bracketing.

other (narrowly) rather than taking the time necessary to consider the two compound lotteries jointly (broadly). Subjects who broadly frame the two compound lotteries are capable of abiding by the Invariance Axiom, i.e., all else equal, they will be more likely to choose Lottery B in Compound Lottery 1 and Lottery A in Compound Lottery 2 because they take the time to compare the two compound lotteries. As Kahneman (2011) points out, in real life broad framing induces *Homo sapiens* to choose high deductibles for insurance policies, eschew choosing extended warranties for the products they purchase, and not regularly check their retirement balances. Broad framing encourages adherence to "risk policies" that lead *Homo sapiens* to make choices with favorable odds in the long run.[6]

TESTING THE INVARIANCE AND DOMINANCE AXIOMS (VERSION 2)

Consider the following two experiments proposed by Tversky and Kahneman (1986):

Experiment 1

The following lottery is described by the percentage of marbles of different colors in each box and the amount of money you win or lose depending upon the color of a randomly drawn marble. Which lottery do you prefer?

Lottery A

Purple 90% chance to win $0
Red 6% chance to win $45,000
Green 1% chance to win $30,000
Blue 1% chance to lose $15,000
Grey 2% chance to lose $15,000

Lottery B

Purple 90% chance to win $0
Red 6% chance to win $45,000
Green 1% chance to win $45,000
Blue 1% chance to lose $10,000
Grey 2% chance to lose $15,000

Experiment 2

The following lottery is described by the percentage of marbles of different colors in each box and the amount of money you win or lose depending upon the color of a randomly drawn marble. Which lottery do you prefer?

Lottery A

6. See Read et al. (1999b) for a seminal discussion on the topic of choice bracketing among *Homo sapiens*.

Purple 90% chance to win $0
Red 6% chance to win $45,000
Green 1% chance to win $30,000
Grey 3% chance to lose $15,000

Lottery B

Purple 90% chance to win $0
Red 7% chance to win $45,000
Green 1% chance to lose $10,000
Grey 2% chance to lose $15,000

Begin by noting that, just as in Version 1 above, Lottery B in Experiment 1 dominates Lottery A. This is because the expected winnings from Lottery B are greater than those from Lottery A, and the expected losses from Lottery B are less than those from Lottery A. *Homo economicus* chooses Lottery B and thus again does not violate the Dominance Axiom.

Next, note that Experiment 2 is effectively identical to Experiment 1. Specifically, Lotteries A in both experiments offer the same percentages of winning $0, $45,000, and $30,000, respectively, and the same percentage of losing $15,000. Lotteries B similarly offer the same percentages of winning $0 and $45,000, respectively, and the same percentages of losing $15,000 and $10,000, respectively.

Since *Homo economicus* chooses Lottery B in Experiment 1 via the Dominance Axiom, in Experiment 2 she will also choose Lottery B, thus abiding by the Invariance Axiom.

What about *Homo sapiens*? Thankfully, none of Tversky and Kahneman's subjects in Experiment 1 chose Lottery A, implying again that *Homo sapiens* also abide by the Dominance Axiom. However, in Experiment 2, a slight majority of subjects chose Lottery A. This is another demonstration of preference reversal for these *Homo sapiens*, which, as we now know well, is a violation of the Invariance Axiom.

TESTING THE SUBSTITUTION AXIOM

Consider the following experiments proposed by Kahneman and Tversky (1979):

Experiment 1

Choose between lotteries **A** and **B**:

A 45% chance to win $6,000
B 90% chance to win $3,000

Experiment 2

Choose between lotteries **A** and **B**:

A 0.1% chance to win $6,000
B 0.2% chance to win $3,000

Homo economicus notices two things about Experiments 1 and 2. First, in each experiment, the two lotteries have identical expected payoffs. In Experiment 1, the expected payoff is $2,700 for both Lotteries A (0.45 x $6,000 = $2,700) and B (0.9 x $3,000 = $2,700), and in Experiment 2, the expected payoff is $6 for both Lotteries A (0.001 x $6,000 = $6) and B (0.002 x $3,000 = $6). Thus, we would expect a sample of *Homo economicus* to split roughly 50%-50% in choosing between Lotteries A and B in Experiment 1 and 50%-50% in choosing between Lotteries A and B in Experiment 2. Second, being the omniscient creature that he is, *Homo economicus* also recognizes that the probabilities in Experiment 1 for Lotteries A and B are actually multiplied by a common factor of 0.002 to obtain the corresponding probabilities in Experiment 2 for Lotteries A and B. Thus, *Homo economicus* understands fully the substitution that has occurred here between the two experiments.

It's a different story for *Homo sapiens*. Based upon a sample of 66 students, Kahneman and Tversky found a split of 14%-86% between Lotteries A and B in Experiment 1, and 73%-27% between Lotteries A and B in Experiment 2. In other words, the sample of *Homo sapiens* seems to have understood neither that the expected payoffs for each lottery are equal in each respective lottery, nor that the percentages in Experiment 2 are merely substitutes for the percentages in Experiment 1. This latter miscue is what leads *Homo sapiens* to violate the Substitution Axiom.

Not to diminish the importance of their having violated this axiom, it is worth mentioning a possible explanation for why at least some of Kahneman and Tversky's students did so. Note that Lottery B's 90% of winning in Experiment 1 is near certainty. To the extent that they prefer certain outcomes (i.e., are influenced by a "certainty effect"), we therefore might expect the students to prefer Lottery B in Experiment 1, despite Lottery A's and B's equal expected values. Furthermore, after having participated in Experiment 1, the probabilities of winning in Experiment 2 (i.e., 0.01% and 0.02%) could very well have seemed inconsequential to a number of the students. Hence, because of such low probabilities of winning (i.e., such inconsequentiality), the students could be forgiven for having chose Lottery A in Experiment 2. The stakes are so low in this experiment, why not take a chance on Lottery A?

TESTING THE SURE-THING PRINCIPLE

Consider the following experiments proposed by Tversky and Shafir (1992):

Experiment 1

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and rundown, and you are not sure that you passed the exam. If you failed you will have to take the exam again in a couple of months—after the semester break. You now have an opportunity to buy a very attractive 5-day vacation package to the Bahamas at an exceptionally low price. The special offer expires tomorrow, while the exam grade will not be available until the day after tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a $150 non-refundable fee to retain the right to buy the vacation package at the same low price the day after tomorrow—after you learn whether you passed the exam.

Experiment 2

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and rundown, and you find out that you passed the exam. You now have an opportunity to buy a very attractive 5-day vacation package to the Bahamas at an exceptionally low price. The special offer expires tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a $150 non-refundable fee to retain the right to buy the vacation package at the same low price the day after tomorrow.

Experiment 3

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and rundown, and you find out that you failed the exam. You now have an opportunity to buy a very attractive 5-day vacation package to the Bahamas at an exceptionally low price. The special offer expires tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a $150 non-refundable fee to retain the right to buy the vacation package at the same low price the day after tomorrow.

These experiments do not provide a clear context within which to test *Homo economicus* vs. *Homo sapiens* (which is nice for a change, given that *Homo sapiens* have thus far paled in comparison to *Homo economicus* in terms of not violating our cherished rationality axioms). Rather, because Experiments 2 and 3 are "sure things" in the sense that the outcome of the exam is known before the decision is made about whether to purchase the vacation package, and Experiment 1 is an "unsure thing" given that the result of the exam is unknown prior to making the decision, we would expect that if the percentages of those participants choosing A, B, and C in Experiment 2 are roughly equal to their corresponding percentages in Experiment 3, then these same percentages should in turn roughly equal those in Experiment 1. In other words, the percentage of participants in Experiment 1 choosing A should roughly equal the percentage of participants in Experiment 2 choosing A, which should roughly equal the percentage of participants in Experiment 3 choosing A, and so on for choices B and C across the experiments. In other words, we would expect that the uncertainty embodied in Experiment 1 should not cause its results to noticeably deviate from the results in Experiments 2 and 3.

Using different samples of roughly 70 students per experiment, Tversky and Shafir report the following results (in percentages):

Experiment		Choice	
	A	B	C
1	32	7	61
2	54	16	30
3	57	12	31

Note that the percentages across choices A – C are roughly the same for "sure-thing" Experiments 2 and 3. But the percentages deviate quite considerably from those for "unsure-thing" Experiment 1. Thus, Tversky and Shafir claim that their groups of *Homo sapiens* violate the Sure-Thing Principle.

STUDY QUESTIONS

1. Can you design a laboratory experiment to test the axiom that was not considered in this chapter—the Independence Axiom?

2. In Testing the Invariance and Dominance Axioms (Version 1), it is clear that broad framing leads to better outcomes than narrow framing. Can you think of a situation in real life where narrow framing could lead to a better outcome than broad framing?

3. Which biases discussed in Chapter 2 are most likely to be avoided through broad framing? Explain.

4. Can you think of three deficiencies associated with the laboratory experiments (discussed in this chapter) that were conducted with the researchers' own students?

5. State in words why a violation of the Transitivity Axiom introduced in Chapter 3 also implies what we are calling a "preference reversal" in this chapter (e.g., see Testing the Invariance Axiom (Version 3)).

6. Can you design a simpler experiment to test for the Sure-Thing Principle than the one presented in this chapter?

Media Attributions

- Table 1 (Chapter 5) © Arthur Caplan is licensed under a CC BY (Attribution) license

CHAPTER 6.

LABORATORY EXPERIMENTS: ADDITIONAL DIFFERENCES BETWEEN *HOMO ECONOMICUS* AND *HOMO SAPIENS*

As mentioned previously, this chapter presents additional laboratory experiments designed to test the implications of the theories advanced in Chapter 4. Here, we learn about the classic advances made by behavioral economists and the main concepts underscored by Prospect Theory; concepts such as mental accounting, Ambiguity and Competency Effects, fairness, regret and blame, as well as loss aversion, reference dependence, and the Endowment Effect.

MENTAL ACCOUNTING (VERSION 1)

Consider the following two experiments proposed by Kahneman and Tversky (1984):

Experiment 1

Imagine that you have decided to see a new movie at your local cinema. You went online ahead of time, purchased a ticket for $10, and then printed the ticket to take with you to the cinema. As you enter the cinema, you discover that you have lost the ticket. The ticket cannot be recovered.

Would you pay $10 at the box office for another ticket?

Experiment 2

Imagine that you have decided to see a new movie at your local cinema, which costs $10 for a ticket. As you approach the box office to pay for a ticket, you discover that you have lost $10.

Would you still pay $10 for a ticket to the movie?

Homo economicus would recognize that, regardless of whether he had the $10 ticket in hand but lost it or lost $10 in cash beforehand, once at the cinema the $10 reduction in his income is what's known as a "sunk cost." He would therefore ignore this cost—completely put it out of his mind—and instead answer the question, "Is watching this movie worth $10 to me at this moment?" If the answer is "yes," then he purchases the ticket and watches the movie. If the answer is "no," he heads back home and does not watch the movie. Most importantly, *Homo economicus*' answer to the question is not dependent on whether he lost the ticket itself (as in Experiment 1) or the cash (as in Experiment 2) (i.e., *Homo economicus* would not be guilty of "narrowly framing" his answer on whether it was a ticket or cash that was lost). As a result, we would expect the percentage of *Homo economicus* choosing to pay for another ticket in each experiment to be roughly 50%.

Based on samples of roughly 200 subjects each for two similar experiments, Kahneman and Tversky found that 46% of the subjects in Experiment 1 answered "yes," they would pay $10 at the box office for another ticket, while in Experiment 2, 88% answered "yes." The authors conclude that going to the cinema is normally viewed as a transaction in which the cost of the ticket is exchanged for the experience of seeing the movie. Buying a second ticket increases the cost of seeing the movie to a level that many *Homo sapiens* find unacceptable. In contrast, the loss of cash is not posted to the mental account of the movie, and it affects the purchase of a ticket only by making the individual feel slightly less affluent.[1]

This evidence suggests that *Homo sapiens* is prone to mental accounting while *Homo economicus* is not.

MENTAL ACCOUNTING (VERSION 2)

Consider the following experiments proposed by Kahneman and Tversky (1984):

Experiment 1

Imagine that you are about to purchase a jacket for $70 and a pair of earbuds for $30 (from the same department store). The electronics salesperson informs you that the earbuds you want to buy are on sale for $15 at the other branch of the store, which is a 20-minute drive across town.

Would you make the trip to the other store?

Experiment 2

Imagine that you are about to purchase a jacket for $30 and a pair of earbuds for $70 (from the same department store). The electronics salesperson informs you that the earbuds you want to buy are on sale for $55 at the other branch of the store, which is a 20-minute drive across town.

Would you make the trip to the other store?

A *Homo economicus* participating in Experiment 1 would recognize the same thing as a *Homo economicus* participating in Experiment 2—he saves $15 by making the trip to the other store. Further, we can say that *Homo economicus* would not distinguish between $15 saved on a relatively cheap vs. expensive pair of earbuds, and thus, we would expect 50% of the *Homo economicus* in each experiment to choose to make the trip to the other store. All else equal, we should expect the full cost of traveling

1. Heath and Soll (1996) present evidence from three similar mental-accounting experiments that reach the same conclusions as Kahneman and Tversky—Homo sapiens tend to be avid mental accountants.

to the other store to exceed the $15 savings for half of the *Homo economicus* who would therefore choose not to make the trip.

When it comes to *Homo sapiens*, Kahneman and Tversky found that 68% of their roughly 100 subjects in Experiment 1 chose to make the trip to the other store, while only 29% of their subjects in Experiment 2 chose to make the trip. This is another example of mental accounting, whereby *Homo sapiens* relate the savings associated with making the trip to a reference point that is determined by the context in which the decision arises. In this case, a larger percentage of *Homo sapiens* interpreted the savings on the cheaper pair of earbuds to be worth the trip to the other store. Apparently, saving money on a cheaper pair of earbuds is more valuable than saving the same amount of money on a more expensive pair.

This example relates to what Thaler (1980 and 1985) calls transactional utility, whereby consumers base their purchase decisions on whether they derive value from the belief that they are getting a good deal rather than just the utility derived from the actual item purchased, or what Thaler calls acquisitional utility. The extent to which purchasing decisions are driven by transactional utility helps explain why stores consistently mark certain products as being "on sale." For example, consumers are more likely to buy a product marked "on sale" for $4 when it regularly sells for $6 than they are to buy the same product simply marked as $4. The product may be for sale at the same price, but consumers feel they are getting a better deal when it is "on sale" than when they pay a regular price.

Could there be a better explanation for these experimental results, perhaps something more conclusive to say about a possible reference point? Looking again at the two experiments, we see that the price differential in Experiment 1 results in a considerably larger percentage gain in savings than the differential in Experiment 2 (specifically, the 50% savings in Experiment 1 is more than double the approximately 20% savings in Experiment 2). Thus, to the extent that Kahneman and Tversky's subjects based their decisions in this context on percentage savings rather than the actual dollar amount saved, we would expect deviations from *Homo economicus*' decision. It may be that several of the experiment's subjects behaved as if their reference point for making a choice was percentage rather than actual savings.

MENTAL ACCOUNTING (VERSION 3)

Consider the following experiment:

> Two avid sports fans plan to travel 25 miles to see their favorite basketball team—the Utah Jazz—play a game at Vivint Arena. One of the fans, Patricia, already paid for her ticket. The other, Peter, was on his way to purchase a ticket when he got one free from a friend. A huge blizzard is announced for the night of the game. Which statement best describes the likely outcome of this situation?
>
> **A** Patricia is most likely to brave the blizzard to see the game.
> **B** Peter is most likely to brave the blizzard to see the game.
> **C** Both are equally likely to brave the blizzard to see the game.

Because both Patricia and Peter have tickets to see the game, *Homo economicus* rationalizes that each will simply weigh the expected benefit of braving the blizzard to see the game (which is the psychic joy associated with watching the Jazz compete against the opposing team at the Vivint Center) against the expected cost (the danger of venturing out into the blizzard). Not that this is really germane to the issue at hand (it's more a nod to the nitpicky among you), but the costs of parking and gasoline to power their vehicles, as well as prorated auto insurance and depreciation of their vehicles and their opportunity costs of the time spent traveling to and from the Vivint Center and watching the game itself are not added to their expected costs because *Homo economicus* correctly assumes that Patricia and Peter already accounted for those costs when the tickets were purchased and accepted for free, respectively. Regardless, *Homo economicus* would choose statement C. Those of you who chose statement A suffer from the Sunk Cost Effect. You are mental accountants. There's no known explanation for those of you who chose statement B.

Gourville and Soman (1998) test a slight variation of this experiment:

> One year ago, Mr. Adams paid $40 cash for a ticket to a basketball game to be played later this week. Yesterday, Mr. Baker paid $40 cash for a ticket to the same game. Both men have equally anticipated this game. On the day of the game, there is a snowstorm. Who is more likely to brave the storm and attend the game, Mr. Adams who paid for his ticket long ago, or Mr. Baker who just recently incurred the $40 expense?
>
> **A** Mr. Baker is most likely to brave the snowstorm to see the game.
> **B** Mr. Adams is most likely to brave the snowstorm to see the game.
> **C** Both men are equally likely to brave the snowstorm to see the game.

Similar to the previous experiment with Patricia and Peter, we would expect *Homo economicus* to pick statement C. When it comes to braving the same snowstorm, it shouldn't matter who paid when. As Gourville and Soman explain, the timing of Mr. Adam's and Mr. Baker's ticket purchases should have no impact on their decision to attend the basketball game. Each should accept that the $40 already spent is a sunk cost and base his decision to go to the game solely upon the perceived incremental costs and benefits of going. Facing the same incremental costs and benefits, Mr. Adam's and Mr. Baker's likelihood of attending the game should be equal.

But when it comes to *Homo sapiens*, all bets are off. Gourville and Soman hypothesize that, in keeping with the prevailing wisdom, both Mr. Adams and Mr. Baker will be prone to the Sunk Cost Effect on purchasing their respective tickets. However, Mr. Adams will have gradually adapted to his "upstream ticket purchase" over the year, thus diminishing the Sunk Cost Effect on his decision of whether to attend the game. The authors call this a Payment Depreciation Effect. To the contrary, Mr. Baker—who has had little time before the game to adapt to the cost of his ticket—will perceive the full Sunk Cost Effect of his purchase when deciding whether to attend. Consequently, Gourville and Soman predict that Mr. Baker will therefore be more likely to attend the game.

The authors test their hypothesis in a series of field experiments with individuals at a shopping mall and laboratory experiments with students. They find support for the Payment Depreciation Effect in a variety of contexts. For example, in a laboratory experiment with over 40 students at the University of Colorado, Gourville and Soman presented the subjects with three tasks spread over three weeks. The first two tasks entailed a short and a long survey, each involving a subject's evaluation of popular soft drinks. The short survey was designed to require minimal effort and was expected to take approximately five minutes to complete. Thus, in completing the short survey the subject experienced virtually no cost. In contrast, the long survey was designed to require considerable effort and was expected to take approximately 30 minutes to complete. Therefore, completing the long survey exacted a high cost on the subject.

These first two tasks were separated in time by three weeks with the order of the two surveys randomized across subjects—approximately half of the subjects first completed the short survey, then experienced the three-week delay, and then completed the long survey ("no-delay condition" in terms of having incurred the high cost), while the remaining subjects first completed the long survey, then experienced the three-week delay, and then completed the short survey ("delay condition" in terms of having incurred the high cost). Upon completing the second survey, Gourville and Soman paid each subject $7 and then presented the subject with a third and final task—an ostensibly unrelated exercise in which the subject faced a real-money gamble.

The subjects were told they could bet up to $2, in increments of $0.25, on a single roll of a pair of dice. They were told that if they rolled a seven or greater, they would double their bet, but if they rolled a number less than seven, they would lose their bet. They were asked to indicate the amount they were willing to gamble, after which they were asked to roll the dice. Based on the outcome of that roll, the amount they had indicated was either added to or subtracted from their $7 payment they had earlier received for completing the second survey. In keeping with their hypothesis about the Payment Depreciation Effect, the authors expected the subjects in the delay condition to experience less of a Sunk Cost Effect and, therefore, be more likely to gamble more of their $7 payment than subjects in the no-delay condition.

The authors ultimately found that larger numbers of subjects experiencing the delay condition wagered more of their compensation payment on the gamble than subjects experiencing the no-delay condition. The majority of the subjects experiencing the delay condition wagered between $1 and $2, while slightly more than half of the subjects experiencing the no-delay condition wagered between $0 and $1.

MENTAL ACCOUNTING (VERSION 4)

Consider the following experiments conducted by Prelec and Loewenstein (1998):

Experiment 1

Imagine that you are planning a one-week vacation to the Caribbean that will occur six months from now. The vacation will cost a total of $1,200. Which of the following two options would you choose for financing the vacation?

A Six monthly payments of $200 each during the six months before the vacation.

B Six monthly payments of $200 each during the six months beginning after you return from the vacation.

Experiment 2

Imagine that, six months from now, you are planning to purchase a clothes washer and dryer for your new home. The two machines together will cost $1,200. Which of the following two options would you choose for financing the two machines?

A Six monthly payments of $200 each during the six months before the machines arrive.

B Six monthly payments of $200 each during the six months beginning after the machines arrive.

When presented with Experiment 1, the authors found that 60% of the roughly 90 participants (visitors to the Phipps Conservatory in Pittsburgh) opted for the earlier payments described in option A despite Prelec and Loewenstein's estimate of an implicit interest penalty equaling approximately $50 per participant. However, in Experiment 2, 84% of the same subjects preferred to postpone payments until the washer and dryer arrive (and thus begin paying each month for the next six months after delivery). Thus, Prelec and Loewenstein found that *Homo sapiens* prefer to decouple payments for durable goods such as washing machines (and thus prorate their payments as their benefits from using the goods occur over time), but not necessarily for goods such as vacations, whose benefits do not extend over time (fond memories of the experience notwithstanding).[2] This suggests that *Homo sapiens* fine-tune their mental accounts according to the type of good in question.

DISCOUNTING

Loewenstein and Prelec (1992) compared the outcomes of two experiments to understand how *Homo sapiens* fine-tune their time discounting behavior. The two experiments are as follows:

Experiment 1

Suppose you bought a TV on a special installment plan. The plan calls for two payments; one this

2. While Prelec and Loewenstein's research focuses on how consumers make purchasing and payment decisions, Shefrin and Thaler (1988) consider how consumers essentially decouple their sources of wealth to afford their purchases. Shefrin and Thaler's Behavioral Life-Cycle Hypothesis suggests that consumers either mentally or physically (or both) classify their wealth into one of three accounts: current income, current wealth, and future wealth. Current income is a consumer's account that is meant to be spent in the shorter term (e.g., paycheck-to-paycheck). Current wealth is meant to accumulate over time to enable the purchase of more expensive items than would normally be covered paycheck-to-paycheck. And future wealth is money saved for future consumption (e.g., retirement savings).

week and one in six months. Which of the following two options would you choose for financing the TV?

A An initial payment of $160 and a later payment of $110.

B An initial payment of $115 and a latter payment of $160.

Experiment 2

Suppose you bought a TV on a special installment plan. The plan calls for two payments of $200; one this week and one in six months. Happily, however, the company has announced a sale that applies retroactively to your purchase. Which of the following two options would you choose for financing the TV?

A A rebate of $40 on the initial payment and a rebate of $90 on the later payment.

B A rebate of $85 on the initial payment and a rebate of $40 on the later payment.

As the authors point out, since options A and options B are the same across Experiments 1 and 2 in terms of payment levels and delivery times, we would expect to see no systemic differences in responses from *Homo economicus* participants across the two experiments. When it comes to *Homo sapiens*, however, Loewenstein and Prelec find that a higher percentage of the 85 students who participated in the two experiments opted for the lower discount (i.e., greater earlier payment) represented by option A of Experiment 1—where the question is framed as a loss (i.e., a payment)—rather than as a gain (i.e., a rebate) as in option A of Experiment 2. Specifically, 54% of the subjects participating in Experiment 1 stated a preference for option A over B. However, only 33% of the subjects preferred option A over B in Experiment 2.

To explain these results, the authors argue that in Experiment 1 subjects discount future payments less (i.e., future payments loom larger in a subject's mind), which leads subjects to base their choice upon the size of the total payment—option A's total payment of $270 is less than option B's total payment of $275. In Experiment 2, however, the outcomes are framed as gains and are smaller in magnitude, both of which contribute to relatively high discounting of the rebates received in the future, leading to a preference for option B which offers a greater initial rebate. Hence, when it comes to discounting the future, *Homo sapiens*' choices are influenced by context, in this case, whether options are framed as future payments due or future rebates to be received. As we well know, *Homo economicus* is not swayed by this type of framing.[3]

OVERWEIGHTING IMPROBABLE EVENTS

Consider the following two experiments proposed by Kahneman (2011):

3. See Loewenstein (1987) for additional experiments on issues pertaining to discounting.

Experiment 1

Choose between lotteries **A** and **B**:

A 0.001% chance to win $50,000

B win $5,000 for certain

Experiment 2

Choose between lotteries **A** and **B**:

A 0.001% chance to lose $50,000

B lose $5,000 for certain

Ok, you know the drill. Let's start with what *Homo economicus* would do here. She would calculate the expected payoffs from the two lotteries in each experiment and choose accordingly. Hence in Experiment 1, Lottery B would certainly be chosen since ($5,000 x 1) = $5,000 > ($50,000 x 0.00001) = $0.5, and in Experiment 2, Lottery A would be chosen since (-$50,000 x 0.00001) = -$0.5 > (-$5,000 x 1) = -$5,000. Kahneman reports that in an experiment with roughly 100 subjects, 72% chose Lottery A in Experiment 1 and 83% chose Lottery B in Experiment 2—a marked deviation from what we expect of omniscient *Homo economicus*.

The results from Experiment 1 highlight *Homo sapiens*' proclivity to gamble, while the results from Experiment 2 suggest why we tend to purchase insurance against possible loss. Kahneman and Tversky (1979) go a step further in interpreting these results. They point out that the two experiments demonstrate what they call Possibility and Certainty Effects. On the one hand, lotteries consisting of extremely low probabilities of winning (e.g., 0% – 5%) are still enough to tempt individuals with the possibility of winnings (the Possibility Effect, as demonstrated in Experiment 1). On the other hand, lotteries consisting of extremely high probabilities of losing (e.g., 95% – 100%) are enough to scare individuals into choosing alternatives with lower probabilities, even if those alternatives are associated with high losses (the Certainty Effect, as demonstrated in Experiment 2). As part of their Prospect Theory (discussed earlier in Chapter 4), Kahneman and Tversky point out that *Homo sapiens* are prone to interpret these extremes as depicting discrete shifts in a lottery's odds, and thus, the lottery's expected payoff is essentially ignored. In our minds, we *Homo sapiens* tend to overweight small probabilities (i.e., improbable events) and underweight moderate and high probabilities (i.e., more likely events).

Based upon their accumulated laboratory experiences over the years, Tversky and Kahneman (1992) eventually proposed decision weights, as shown in the table below (Kahneman, 2011). Recall from Chapter 4 that these weights are derived from a decision-weight formula that effectively transforms a lottery's objective probabilities of the different outcomes into their corresponding subjective probabilities, or decision weights.

Probabilities	0	1	2	5	10	20	50	80	90	95	98	99	100
Decision Weights	0	5.5	8.1	13.2	18.6	26.1	42.1	60.1	71.2	79.3	87.1	91.2	100

The row of Probabilities represents objective probabilities that could conceivably define a given lottery. The row of Decision Weights represents subjective probabilities that Kahneman and Tversky (1992) suggest we *Homo sapiens* tend to subconsciously substitute for the objective ones. Note that for lower objective probabilities, the corresponding decision weights are larger (reflecting our penchant for overweighting of improbable events), while for mid- to higher-probabilities, the weights are smaller (reflecting our penchant for underweighting more likely events).

AMBIGUITY AND COMPETENCY EFFECTS

Consider the following experiments conducted by Heath and Tversky (1991):

Experiment 1

Choose between lotteries **A** and **B**:

A A jar contains 50 red marbles and 50 green marbles. Blindly draw a marble and guess its color. If your guess is correct, you win $100.

B A jar contains 100 red and green marbles in unknown proportion. Blindly draw a marble and guess its color. If your guess is correct, you win $100.

Experiment 2

Choose between lotteries **A** and **B**:

A A stock is selected at random from the New York Stock Exchange. You guess whether its price will have gone up or down at close tomorrow. If your guess is correct, you win $100.

B A stock is selected at random from the New York Stock Exchange. You guess whether its price went up or down at close yesterday. You cannot check any news sources. If your guess is correct, you win $100.

Given that *Homo economicus* has no particular color preference, he will be indifferent between Lotteries A and B in Experiment 1. This is because the added information provided in Lottery A is superficial in terms of affecting the outcome associated with drawing a marble from the jar. Thus, given any sample of *Homo economicus*, we would expect 50% to choose Lottery A and 50% to choose Lottery B. In his run of this experiment, Ellsberg (1961) found a larger percentage of *Homo* *sapiens* chose Lottery A than Lottery B in what came to be known as the Ellsberg Paradox. Why? Ellsberg postulates that, although superfluous, the added information provided in Lottery A gives the impression of Lottery A seeming less ambiguous than Lottery B. Thus, *Homo sapiens* tend to submit to what has also come to be known as an Ambiguity Effect.[4]

Similarly, in Experiment 2, *Homo economicus* recognizes that Lotteries A and B are effectively identical. Given the prohibition in Lottery B on one's ability to check the newspaper or check online, the outcome common to each lottery—the change in a randomly chosen stock's price—is unaffected by whether the price change occurred yesterday or tomorrow. Thus, given any sample of *Homo economicus*, we would again expect 50% to choose Lottery A and 50% to choose Lottery B. Lo and behold, in their experiment with roughly 200 subjects, Heath and Tversky found that 67% chose Lottery A. Why?

The authors postulate that *Homo sapiens* naturally prefer appearing competent or, alternatively stated, prefer not to appear incompetent. By choosing a lottery where the outcome is still to be determined in the future, we perceive ourselves as running less risk of appearing incompetent if the worst outcome occurs (e.g., we guess that the stock price will rise when instead it falls) than if we choose the worst outcome of a lottery where, technically speaking, the outcome has already occurred. In other words, *Homo sapiens* tend to submit to what has come to be known as a Competency Effect.[5]

THE DECOY EFFECT

Consider the following experiments proposed by Ariely (2008):

4. In a more recent laboratory experiment, Halevy (2007) presented subjects with four different boxes, each containing some configuration of red and black balls. The subjects were tasked with imagining themselves choosing a box and then guessing which color ball would be randomly chosen from the box. If she guessed the color correctly, the subject would win $2. Box 1 contained five red balls and five black balls. Box 2 contained 10 balls total, but it was unknown how many were red and black. For Box 3 a number between zero and 10 would first be randomly chosen to determine the number of red balls (the remainder would be black balls). And for Box 4 a fair coin would be tossed to determine whether all the balls in the box would be black or all red. Subjects were given the chance to sell bets on each respective box by announcing the amounts they were willing to accept (WTA) to forego the bets. Since (1) $2 was the most a subject could win if they were to actually choose a box and then correctly guess the color of a ball chosen from the box, (2) the subject received $0 if she guessed incorrectly, and (3) there was effectively a 50% chance of guessing the color correctly regardless of the box chosen, a risk-neutral *Homo economicus* subject would accept $1 for each of the boxes. If Halevy's subjects succumbed to the Ellsberg Paradox, then their WTA would be higher for Box 3 or 4 than for Box 2 given the greater amount of ambiguity associated with the former two boxes. The author found the opposite—all but one subject stated a higher WTA for Box 2 than either Box 3 or 4.
5. Competency and Ambiguity effects play out in another interesting context. Thaler et al. (1997) conducted an experiment where subjects were instructed to choose what percentage of their funds to invest in a relatively risky asset (larger expected return but with more risk) versus a relatively safe asset (lower expected return but with less risk). The experiment consisted of two treatments. The first treatment required subjects to make their respective decisions every period about their portfolio allocation; the second treatment allowed subjects to make portfolio-allocation decisions only once per eight periods. Subjects in the first treatment placed more than 50% of their portfolio in the safe asset over the course of the experiment. In contrast, subjects in the second treatment placed only between 30% and 40% of their portfolios in the safe asset. Thus, by simply increasing the frequency with which subjects were required to evaluate the performance of their portfolios, the researchers found that subjects became more sensitive to risk. Could this be evidence of a Competency Effect, or is it more an instance of "ignorance is bliss"? Either way, Thaler et al. (1997) identify this quirky behavior as evidence of narrowly bracketed investment decision-making.

Experiment 1

Suppose *The Economist* magazine runs the following advertisement announcing a new deal on annual subscription rates. New subscribers have three options.

1. Internet-only subscription for $59. Includes online access to all articles published in *The Economist* since 2010.
2. Print-only subscription for $125. Includes print copies of *The Economist* mail-delivered for the current year.
3. Print-and-internet subscription for $125. Includes online access to all articles published in The Economist since 2010 and print copies of *The Economist* mail-delivered for the current year.

You have decided that you would like to begin reading *The Economist*, and thus must choose one of the three options. Which option would you choose?

Experiment 2

Suppose *The Economist* magazine runs the following advertisement announcing a new deal on annual subscription rates. New subscribers have two options.

1. Internet-only subscription for $59. Includes online access to all articles published in *The Economist* since 2010.
2. Print-and-internet subscription for $125. Includes online access to all articles published in The Economist since 2010 and print copies of *The Economist* mail-delivered for the current year.

You have decided that you would like to begin reading *The Economist*, and thus must choose one of the two options. Which option would you choose?

Clearly, all *Homo economicus* participating in these two experiments who wish to have print versions of *The Economist* will choose the print-and-internet subscription for $125. It doesn't matter that the print-only option is missing in Experiment 2. *Homo economicus* who wish to have print copies of the magazine would never choose that option as long as the print-and-internet option is available. They get more for their money with the print-and-internet option and would therefore never pass it up.

This is what Ariely found for Experiment 1 when 100 of his students at MIT's Sloan School of Management were presented with the three options. Sixteen students chose the internet-only option and 84 chose the print-and-internet option. No students chose the print-only option. However, when a different group of 100 students participated in Experiment 2, 68 chose the internet-only option and only 32 chose the print-and-internet option. What happened?

As Ariely describes it, this result is the economic equivalent of the theory of relativity; relativity that exposes predictably irrational choice behavior among *Homo sapiens*. In this particular case, the mere presence of the print-only subscription in Experiment 1 served as a "decoy" that sent 84 of the students to the print-and-internet option. The absence of the decoy in Experiment 2 led the students in that experiment to choose differently. Only 32 students chose the print-and-internet subscription.

This Decoy Effect is a special case of economic relativity. And as Ariely points out, these types of effects mirror the way *Homo sapiens*' mind is wired—we are prone to consider things around us in a relative sense. This holds true not only for physical things, such as toasters, bicycles, puppies, restaurant entrées, and spouses but also for experiences such as vacations, educational options, emotions, attitudes, and points of view. *Homo economicus*, on the other hand, thinks and acts in a world of absolutism.

THE ZERO-PRICE EFFECT

Consider the following two experiments proposed by Shampanier et al. (2007):

Experiment 1

Suppose you are given $1 to participate in this experiment. You are presented with the choice of purchasing a Lindt truffle for 75 cents or a Hershey's Kiss for 25 cents. You can choose one or the other, or neither. What will you choose to do?

Experiment 2

Suppose you are given $1 to participate in this experiment. You are presented with the choice of purchasing a Lindt truffle for 50 cents or getting a Hershey's Kiss for free. You can choose one or the other, or neither. What will you choose to do?

Despite having no prior information about the preferences of any given *Homo economicus* for Lindt truffles vs. Hershey's kisses, we should nevertheless expect that if 50, 25, and 25 individuals out of a sample of 100 *Homo economicus*, respectively, were to select the Lindt truffle, Hershey's kiss, and neither in Experiment 1, then roughly the same respective numbers will be selected by *Homo economicus* in Experiment 2. Why? Because the difference in prices between the Lindt truffle and Hershey's kiss is the same in each experiment (50 cents). Rationally speaking, *Homo economicus* interprets the two experiments as offering the same choice.

To see this, suppose *Homo economicus* Harry estimates the amount of utility he expects to get from the truffle and the kiss (suppose it's 150 utils and 50 utils, respectively) and then subtracts the disutility he gets from paying for each. Without loss of generality, suppose each cent paid is a dis-util. In Experiment 1, this means that Harry would receive a net utility (or, net benefit) of 150 – 75 = 75 utils by choosing the truffle, and 50 – 25 = 25 utils by choosing the Hershey's kiss. Harry therefore gains 75 – 25 = 50 utils by choosing the truffle. In Experiment 2, the net benefit from choosing the truffle is 150 – 50 = 100 utils and the net benefit from choosing the kiss is 50 – 0 = 50 utils. Again, Harry gains 100 – 50 = 50 utils by choosing the truffle. As far as Harry is concerned, the choices in each experiment are identical.

Not so *Homo sapiens*. Shampanier et al. ran similar experiments with roughly 400 students at the

MIT campus and obtained some surprising results. Shampanier et al.'s prices in Experiment 1 were 15 cents for the truffle and 1 cent for the kiss, and in Experiment 2, the prices were 14 cents for the truffle and 0 for the kiss (thus, the choices in each experiment are again identical). To test just how strong the pull of a free Hershey's kiss might be, the authors had a subgroup of the 400 students instead choose between a 10-cent truffle and a free Hershey's kiss in Experiment 2.

Shampanier et al. found that 36% of the students chose the Lindt truffle, 14% the Hershey's kiss, and 50% chose neither in Experiment 1. However, in Experiment 2, only 19% of the students chose the truffle and 42% chose the free kiss. The percentage of students choosing the free kiss was roughly the same (40%) in the version of Experiment 2 where the price of the truffle dropped to 10 cents rather than 14 cents. This "pull" of the free Hershey's kiss is what Shampanier et al. call the Zero-Price Effect. Apparently, the difference between 15 cents, on the one hand, and 14 or 10 cents on the other is small. But the difference between 1 cent and zero is huge. What gives?

Ariely (2008) puts it this way:

"Most transactions have an upside and a downside, but when something is FREE! we forget the downside. FREE! gives us such an emotional charge that we perceive what is being offered as immensely more valuable than it really is. Why? … it's because humans are intrinsically afraid of loss. The real allure of FREE! is tied to this fear. There's no visible possibility of loss when we choose a FREE! item (it's free). But suppose we choose the item that's not free. Uh-oh, now there's a risk of having made a poor decision—the possibility of a loss. And so, given the choice, we go for what is free" (p. 60).

Hmmm. Sounds like a case of loss aversion.[6]

VIVIDNESS OF PROBABILITY

Consider the following experiments conducted by Kahneman (2011):

Experiment 1

Suppose you are a psychiatrist at a psychiatric hospital. You are in charge of evaluating whether it is safe to discharge Mr. Thomas from the hospital. Mr. Thomas has a history of violence. You have received the following assessment from a criminal expert concerning the risk associated with releasing Mr. Thomas from the hospital:

"Patients similar to Mr. Thomas are estimated to have a 10% probability of committing an act of violence against others during the first several months after discharge."

Will you deny Mr. Thomas' discharge?

6. But that's not all, folks! Ariely et al. (2018) report on experiments where the Zero-Price Effect works in the opposite direction. Lowering the price to zero (in this case of Starburst Fruit Chews) actually leads to a net decrease in the total amount demanded in the market. This occurs when the limit on the total supply of the good in question is either known with certainty or perceived by the consumer, in either case, the good effectively becomes a shared resource. The authors conclude that participants in these experiments applied a simple social-norm rule to their choice behavior. Once the candy was priced at zero, they chose to sacrifice their own desires for the benefit of others. As Ariely (2008) points out, these results suggest why, when we are dining out with friends, taking the last slice of pizza feels so wrong.

Experiment 2

Suppose you are a psychiatrist at a psychiatric hospital. You are in charge of evaluating whether it is safe to discharge Mr. Thomas from the hospital. Mr. Thomas has a history of violence. You have received the following assessment from a criminal expert concerning the risk associated with releasing Mr. Thomas from the hospital:

"Of every 100 patients similar to Mr. Thomas, 10 are estimated to commit an act of violence against others during the first several months after discharge."

Will you deny Mr. Thomas' discharge?

Since the statistics provided in both experiments (blue font) are identical, we would expect the same sample percentages of *Homo economicus* across Experiments 1 and 2 to answer "yes" and "no"—50% and 50%, respectively. Such is not the case with samples of *Homo sapiens* (surprise, surprise). Kahneman reports that in his experiments, only 21% answered "yes" in Experiment 1, while 41% answered "yes" in Experiment 2. Kahneman speculates that fewer subjects answered "yes" in Experiment 1 because of the vividness of the probability provided in Experiment 1, as opposed to the relatively undramatic (for lack of a better word) number provided in Experiment 2. For whatever reason, 10% conjures more of an impact in our minds than 10 out of 100. Who knew?

ENVY AND GUILT (OR INEQUALITY AVERSION, OR FAIRNESS)***

Since we have been exploring effects driven by human emotion (e.g., perceived ambiguity and competency) this is as good a place as any to investigate what behavioral economists have to say about the emotions envy and guilt. There is no experiment here, just some economic conceptualizing. In Chapter 5, we will explore how these emotions manifest themselves as fairness in behavioral games. Here, we take a little detour and investigate what envy and guilt actually look like in the context of a standard neoclassical framework.[7]

Recall from Chapter 2 that the typical *Homo economicus* utility function looks something like,

$$u(x_i) = \sqrt{x_i},$$

where x_i represents individual i's wealth level (yes, in Chapter 2 w_i was used to represent i's wealth level—we make the notational change here to be consistent with the ensuing discussion).

To represent the potential effects of envy and guilt on *Homo sapiens* utility function, Fehr and Schmidt (1999) propose an alternative specification,

$$u(x_i, x_j) = x_i - \alpha_i \ max\{x_j - x_i, 0\} - \beta_i \ max\{x_i - x_j, 0\},$$

where x_j represents another individual j's wealth level, $i \neq j, 0 \leq \alpha_i < 0.5$ represents individual

7. I almost said, "....in the context of the rational-choice world of Homo economicus." However, I chose not to for the simple, albeit technical, reason that Homo economicus is, by definition, devoid of human emotions such as envy and guilt. The model that we present here assumes our Homo economicus has complete information. See Cartwright (2018) for a model with incomplete information.

i's marginal disutility from envy, and $0 \leq \beta_i < 1$ represents individual i's marginal disutility from guilt.[8]

Let's unpack Fehr and Schmidt's utility function. To begin, consider the term $-\alpha_i \; max\left\{x_j - x_i, 0\right\}$. This says that to the extent that individual j's wealth exceeds individual i's, individual i's utility decreases by a constant factor of α_i. If instead individual i's wealth exceeds j's, then there is no negative effect on i's utility because then zero is larger than a negative number. Can you guess what human emotion this term is accounting for? Yep, it's envy.

Now consider the term $-\beta_i \; max\left\{x_i - x_j, 0\right\}$. You will note that the order of subtraction in this term is reversed from the previous term's. Now, to the extent that individual i's wealth exceeds individual j's, individual i's utility decreases by a constant factor of β_j. If instead individual i's wealth exceeds j's, then there is no negative effect on i's utility because then, again, zero is larger than a negative number. This term accounts for the human emotion of guilt.

As you know, we economists like to draw graphs (how does that saying go, a graph is worth a thousand words?). So, let's consider what Fehr and Schmidt's (1999) utility function looks like in the form of an indifference curve (recall our introduction to this curve in Chapter 3).

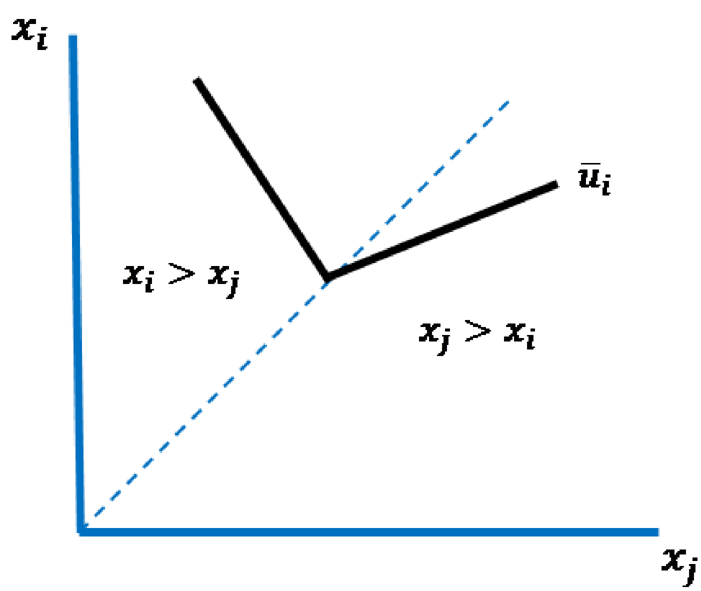
In the case of *Homo sapiens* (who tend to experience the emotions of envy and guilt), the indifference curve for individual i looks different—in some important respects, much different. Most significantly, to account for envy and guilt, individual i's indifference curve must now incorporate the level of some other individual j's wealth, as per Fehr and Schmidt's utility function.

For those of you with a stronger background in economics, you will note that the formula defining an individual's indifference curve solves as,

$$x_i = \frac{\bar{u}}{1-\beta_i} - \frac{\beta_i}{1-\beta_i}x_j \quad \text{if } x_i > x_j\text{, and}$$
$$x_i = \frac{\bar{u}}{1+\alpha_i} + \frac{\alpha_i}{1+\alpha_i}x_j \quad \text{if } x_i < x_j\text{,}$$

which results in an indifference curve for individual i looking like that depicted in Figure 6.1.

Figure 6.1. Indifference Curve for Envious and/or Guilt-Ridden *Homo sapiens*



Note that this indifference curve now depicts the interpersonal tradeoff between the two individuals' wealth levels, x_i and x_j, rather than the intrapersonal tradeoff between the two physical

8. We explain why the upper bound on α_i is 0.5 rather 1 a bit further below.

quantities, x_{i1} and x_{i2}, exhibited by individual i alone. Recall that when wealth levels are such that $x_i > x_j$, individual i is prone to feelings of guilt and is thus confined to the region above the 45° hashed line in the graph. From our formula for the indifference curve above, we see that the slope of the indifference curve's line segment in this region of the graph is equal to $-\beta_i/(1-\beta_i) < 0$. The value of this fraction in turn measures the (constant) rate at which individual i is willing to sacrifice, or transfer, some of her wealth to individual j to assuage her guilt at having more wealth (note that any point above the 45° line indicates that individual i's wealth exceeds individual j's). The larger is β_i, the more steeply sloped is the line segment, implying that individual i feels a greater sense of guilt from her wealth differential with individual j, and thus, is willing to transfer even more of her wealth to j per unit of j's wealth.

You might be wondering what the terminology "per unit of j's wealth" means in this instance. Given that a dollar of wealth to individual i is equal to a dollar of wealth to individual j, doesn't this mean that β_i must equal 0.5, implying that $-\beta_i/(1-\beta_i) = -1$ (i.e., wealth transfers occur on a one-to-one basis in dollars)? In the present context, the answer is "no." In cases where $\beta_i > 0.5 \Longrightarrow -\beta_i/(1-\beta_i) < -1$, individual i feels so guilty that she would willingly transfer more than $1 of her wealth for each $1 of wealth individual j receives. Talk about feeling guilt-ridden!

In contrast, when wealth levels are such that $x_j > x_i$, and therefore, individual i is prone to feelings of envy, individual i is confined to the region below the 45° hashed line in Figure 6.1. Again, appealing to our formula for the indifference curve above, we see that the slope of the indifference curve's line segment in this region of the graph is equal to $\alpha_i/(1-\alpha_i) > 0$. The value of this fraction in turn measures the (constant) rate at which individual i believes his own wealth should be compensated as individual j's wealth increases. Note that larger values of α_i indicate larger feelings of envy. However, the upper bound on α_i is equal to 0.5 (rather than 1, as for β_i) because, as can be seen in the graph, if $\alpha_i > 0.5$, then the indifference curve's associated line segment will automatically cross into the guilt region (where $x_i > x_j$) and thus, be inconsistent with feelings of envy. In other words, individual i's wealth cannot be such that she simultaneously feels envy and guilt about individual j's wealth level.

FAIRNESS IN THE CONTEXT OF FRAMING

Consider the following two experiments conducted by Kahneman (2011):

Experiment 1

A company is making a small profit. It is located in your hometown, which is currently experiencing a recession with substantial unemployment but no price inflation. The company decides to decrease wages and salaries by 7% this year.

On a scale from 1 to 5, with 1 being "very fair" and 5 being "very unfair", how do you rate this action by the company?

Experiment 2

A company is making a small profit. It is located in your hometown, which is currently experiencing a recession with substantial unemployment and price inflation of 12%. The company decides to increase wages and salaries by only 5% this year.

On a scale from 1 to 5, with 1 being "very fair" and 5 being "very unfair," how do you rate this action by the company?

Confronted with these two experiments, *Homo economicus* is not fooled by what's known as "money illusion." She knows the difference between nominal and real changes and the importance of making decisions based upon the latter. As a result, *Homo economicus* recognizes that both experiments ask her to rate the fairness of a company's decision in the face of a recession that is effectively costing the company's employees (i.e., those employees who have been able to retain their jobs with the company during the recession) 7% of their incomes in real terms. In Experiment 1, this 7% loss in real income results from the company's decision to reduce wages and salaries by 7% in the face of no price inflation. In Experiment 2 the loss occurs as a result of the company raising wages and salaries by 5% in the face of 12% price inflation. Thus, in separate samples of *Homo economicus*, we would expect roughly equal percentages of subjects to choose numbers 1 – 5 across the two samples, and for the distribution of percentages to be roughly uniform across the numbers (e.g., 20% choosing 1, 20% choosing 2, etc.).

You probably won't be too surprised to learn that when Kahneman ran this experiment with his students, 62% chose 4 or 5 in Experiment 1 while only 22% chose 4 or 5 in Experiment 2. Clearly, context matters here for *Homo sapiens*. In this case, the context broaches the principle of "dual entitlement" in the mind of *Homo sapiens*, a principle where both the employer and employees are entitled to levels of benefit provided by some "reference transaction."

The context in Experiment 2 is framed as being less unfair than the context in Experiment 1 simply because the company in Experiment 2 appears to be doing something more to protect its employees' wages and salaries in the face of inflation than the company in Experiment 1. After all, the company in Experiment 2 is increasing wages and salaries, not lowering them. Unfortunately, the larger percentage of students rating the company in Experiment 1 as being more unfair have been framed. They suffer from money illusion.

Kahneman et al. (1986a; 1986b) ran additional experiments to measure *Homo sapiens*' penchant for fairness in other contexts. For example, the authors posed the following experiment to approximately 200 adult residents in the Vancouver metropolitan area of British Columbia, Canada:

A football team normally sells some tickets on the day of their games. Recently, interest in the next game has increased significantly, and tickets are in great demand. The team owners can distribute the tickets in one of three ways.

Auction: The tickets are sold to the highest bidders.
Lottery: The tickets are sold to the people whose names are drawn.
Queue: The tickets are sold on a first-come-first-served basis.

Rank these three options from most to least fair.

Being a devoted practitioner of hard, cold economic efficiency, *Homo economicus* would rank the auction first, as this would allocate the tickets to the fans willing to pay the most for them, and queueing last, as this is the most economically wasteful way to allocate resources. To *Homo economicus*, economic efficiency and fairness are one and the same. Unsurprisingly, Kahneman et al.'s participants completely reversed *Homo economicus*' ranking. The great majority of the sample ranked queueing as fairest and the auction as the least fair.

In a second set of experiments, the authors contacted adult residents living in the Vancouver and Toronto metropolitan areas and posed the following two experiments to two separate samples of participants:

Experiment 1

A landlord rents out a single small house to a tenant who is living on a fixed income. A higher rent would mean the tenant would have to move. Other small rental houses are available on the market. The landlord's costs have increased substantially over the past year, and the landlord raises the rent to cover the cost increases when the tenant's lease is due for renewal.

Rate the landlord's decision to raise the tenant's rent as either:

1. Completely fair
2. Acceptable
3. Somewhat unfair
4. Very unfair

Experiment 2

A small photocopying shop has one employee who has worked in the shop for six months and earns $17 per hour. Business continues to be satisfactory, but a factory in the area has closed and unemployment has increased. Other small shops have now hired reliable workers at $14 per hour to perform jobs similar to those done by the photocopy-shop employee. The owner of the shop reduces the employee's wage to $14 per hour.

Rate the shop owner's decision to lower his employee's wage as either:

1. Completely fair
2. Acceptable
3. Somewhat unfair
4. Very unfair

While we would expect *Homo economicus* to rate the decisions by both the landlord in Experiment 1 and the shop owner in Experiment 2 as "completely fair," what about *Homo sapiens*? Given what we (think we) know about their penchant for fairness, we might expect *Homo sapiens* to rate both the landlord and shop owner as "very unfair" or at least "somewhat unfair." Surprisingly, *Homo sapiens* are not that predictable.

While the great majority of participants rate the shop owner as "somewhat unfair" or "very unfair," they view the landlord as being "completely fair" or "acceptable." What's going on here? Kahneman et al. propound two rules governing the fairness judgments made by participants in the respective experiments. In Experiment 1, it is acceptable for the landlord to maintain her profit level at its reference level by raising rent as necessary, even when doing so causes considerable loss or inconvenience for the tenant. To the contrary, in Experiment 2 it is unfair for the shop owner to exploit an increase in his market power to alter his profit from its reference level at the direct expense of an employee.

As these experiments demonstrate, *Homo sapiens* consider fairness to be context-specific.

REGRET AND BLAME

Consider the following experiments proposed by Kahneman (2011):

Experiment 1

Winona very rarely picks up hitchhikers. Yesterday she gave a man a ride and was robbed. Alfred frequently picks up hitchhikers. Yesterday he gave a man a ride and was robbed.

Which of the two—Winona or Alfred—will experience greater regret over the episode?

Experiment 2

Winona almost never picks up hitchhikers. Yesterday she gave a man a ride and was robbed. Alfred frequently picks up hitchhikers. Yesterday he gave a man a ride and was robbed.

Which of the two—Winona or Alfred—will be criticized by others more severely over the episode?

Because *Homo economicus* ignores possible reference points (i.e., is reference independent) and narrowly frames decisions like this, he is not influenced by Winona's and Alfred's past experiences with hitchhikers. Thus, *Homo economicus* would have no reason to assign greater regret to Winona or Alfred in Experiment 1, or more criticism (i.e., blame) to one or the other in Experiment 2. When it comes to assigning greater regret and more blame, *Homo economicus* essentially flips two fair coins—if "heads" then assign greater regret and more blame to Winona, "tails" assign them to Alfred. We would consequently expect respective samples of *Homo economicus* to assign greater regret and more blame 50%-50% to Winona and Alfred.

Kahneman's experiments with his students resulted in 88% assigning greater regret to Winona, and 77% assigning more blame to Alfred. Therefore, appears that *Homo sapiens* are prone to using a social norm as their reference point when deciding how to apportion regret and blame to people like Winona and Alfred. In this case, the norm is "do not pick up hitchhikers." The logic for how this norm *cum* reference point could be driving *Homo sapiens*' choices in these experiments goes something like this:

> *Because Alfred has frequently (and presumably knowingly) flaunted this norm in the past, "he had it coming to him," and thus, "should have seen it coming." Alfred is, therefore, not entitled to feel as much regret as does Winona, who, by contrast, has rarely if ever flaunted the norm. Winona was less likely to see the robbery coming and likely feels greater regret at having transgressed a norm that she has traditionally followed. Because Alfred has traditionally transgressed the norm and had the robbery coming to him, he effectively deserves to be more severely criticized.*

Hopefully, you see that, as compelling as this logic seems, it is misguided. With respect to regret, the question pertains to what we think Winona and Alfred will feel about themselves after having been robbed (think Regret Theory from Chapter 4), not what we think they are entitled to feel. Similarly, regarding the apportionment of blame, we know nothing about the people who will be judging the two victims, in particular, how they tend to judge others' behaviors. In the end, then, this is a case where narrowly framing the situations and avoiding the use of reference points would actually help *Homo sapiens* reach more judicious judgments.

ASYMMETRIC REGRET

Consider the following thought experiment conducted by Kahneman (2011):

> Cheryl owns shares in Company A on the New York Stock Exchange. During the past year, she considered switching to owning Company B's stock, but she decided against it. She now learns that she would have been better off by $25,000 if she had switched to Company B's stock when she had considered doing so.
>
> Wilbur used to own shares in Company B on the New York Stock Exchange. During the past year, he switched to stock in Company A. He now learns that he would have been better off by $25,000 if he had kept his stock in Company B.
>
> Who feels greater regret?

Similar to how she interpreted the human emotion of regret in the previous experiment, *Homo* *economicus* will again ignore potential reference points and narrowly frame her judgment (i.e., she will not abide by the strictures of Regret Theory). In the final analysis, Cheryl and Wilbur both lost $25,000 by choosing to hold stock in Company A. It doesn't matter that Cheryl held onto the stock rather than switching to stock in Company B, or that Wilbur switched from owning stock in Company B to owning stock in Company A. They both lost $25,000, and should therefore both feel the same amount of regret.

Not so with *Homo sapiens*. Kahneman reports that 92% of his subjects assigned a greater sense of regret to Wilbur than to Cheryl. The reference point here pertains to whether an investor decides to switch ownership in a stock or not. The logic goes something like this:

> *When the value of their stock in a given company falls, investors who exhibit inertia and choose not to sell their ownership in that stock beforehand later experience less regret than investors who instead choose to purchase ownership in that stock beforehand. In this case, the proverb "fools rush in where angels fear to tread" implies that when what the fools rushed into costs them money, they should feel more regret than the angels who lost the same amount of money by instead exhibiting more patience.*

Yeah, right. Their propensity for broad framing and reference-dependent decision-making has apparently again led *Homo sapiens* astray ("apparently" being an important word here). Perhaps Kahneman's subjects should have instead flipped fair coins and tried not to reason their ways to answers. Or not. Could there be a more compelling logic for Kahneman's results?

One could argue that, to the extent Kahneman's subjects believed Cheryl and Wilbur are prone to what we previously learned in Chapter 4 is called the Endowment Effect, then the subjects were justified in concluding that Wilbur suffers more regret than Winona. An Endowment Effect occurs when the intrinsic value associated with owning something (e.g., a given commodity) is large enough to induce the owner to unwittingly overprice the commodity in a market setting.[9] Therefore, if Wilbur is susceptible to the Endowment Effect—specifically a retroactive Endowment Effect associated with the shares he once owned in Company B—then it is very likely he is suffering more regret than Winona by having forfeited that endowment. Because she never actually owned stock in Company B, Winona is perforce precluded from the opportunity to suffer an Endowment Effect.

It is important to note that this is not a case of "two wrongs making a right." While it may be "wrong" for Wilbur to exhibit an Endowment Effect, it is not wrong for Kahneman's subjects to assume that he does. Indeed, one could argue that to the extent Kahneman's subjects made this assumption, they made the correct choice—correct not just because they honored the assumption but also because (as will be discussed in Section 3) *Homo sapiens* really are prone to this effect.

THE GENDER GAP

Differences in the choice behaviors between men and women have been the subject of an immense body of research especially with respect to altruistic tendencies (c.f., Brañas-Garza et al., 2018; Eckel and Grossman, 2001; Solnick, 2001) and empathy and forgiveness (Toussaint and Webb, 2005). While the existence of a gender gap is a non-issue for *Homo economicus*—which, after all, can be thought of as a genderless species—gender is generally believed to be a prolific distinguishing feature of the *Homo sapiens* experience.

In an innovative laboratory experiment, Niederle and Vesterlund (2007) measure a gender gap

9. The Endowment Effect is a special case of the Anchoring Effect we encountered earlier (remember the question asking you to guess Gandhi's age when he died?). As pointed out in this book's preamble, the Anchoring Effect is itself a special case of a Framing Effect. So much nomenclature, so little time!

in how men and women respond to competition. Participants solve two sets of math problems, first under a noncompetitive piece-rate scheme and then under a competitive tournament scheme. Participants are then asked to select which of these two compensation schemes they want to have applied to their next set of math problems (i.e., whether they would avoid competition by choosing the piece-rate scheme or compete with other group members in a tournament). This combination of math-problem performance and choice of compensation scheme enabled the authors to determine if men and women of equal performance choose the same compensation scheme.

Each math problem involved adding up five two-digit numbers without the aid of a calculator, but with scratch paper if desired. The numbers were randomly drawn and each problem was presented in the following manner, where the participants are instructed to fill in the sum in the row's blank box:

Add the numbers in the first five boxes as quickly as you can, and write your answer in the sixth blank box.

21	**35**	**48**	**29**	**83**	

Once a participant submits an answer on the computer, a new problem appears jointly with information on whether the former answer was correct. This process continues for five minutes. A record of the participant's number of correct and wrong answers remains on the screen as s/he progresses through the five-minute set of problems. Their final scores are determined by the number of correctly solved problems in the five-minute timespan. Niederle and Vesterlund (2007) chose this approach because it requires both skill and effort and because there was no prior evidence in the extant literature of gender differences in ability on easy math tests. The authors were, therefore, able to rule out performance differences as an explanation for gender differences in the choice of competition level (i.e., choice between the noncompetitive piece-rate and competitive tournament schemes).

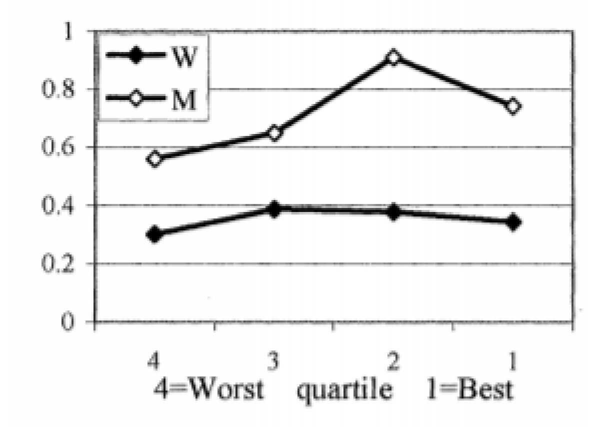
Participants were randomly divided into groups of four, each group seated in a row. Participants were informed that they were grouped with the other people in their row. Each group consisted of two women and two men. Although gender was not discussed at any time, participants could see who the other people in their group were and were thus aware of their group's gender mix. A total of twenty groups participated in the experiment (40 men and 40 women total). Each participant received a $5 "show-up fee" and an additional $7 for successfully completing the experiment.

Participants were instructed to complete four separate sets of math problems (i.e., four separate tasks) and were told that one of these tasks would be randomly chosen for payment after the experiment. While each participant could track their own performance on any given task as the math problems were completed, they were not informed of their relative performance to everyone else in their group until the end of the experiment (i.e., upon conclusion of the fourth task). Under the piece-rate scheme, participants earned $0.50 per correct answer. Under the tournament scheme, the participant who correctly solved the largest number of problems in the group received $2 per correct answer while the other participants received no payment. In case of ties, the winners were chosen randomly from among the high scorers.

Task 1 was presented to each participant under the piece-rate scheme, and Task 2 was presented under the tournament scheme. These initial tasks were meant to serve as baseline measures of each participant's performance under the two schemes. Under Task 3, participants selected whether they wanted to be paid according to the piece-rate or the tournament scheme before engaging in the task. A participant choosing the tournament received $2 per correct answer if her score in Task 3 exceeded that of the other group members in Task 2's tournament they had previously completed. Otherwise, he received no payment. Again, in case of ties, the winners were chosen randomly. Thus, participants choosing to play in a tournament in Task 3 are competing against other participants who had already participated in the two previous tasks, which enabled Niederle and Vesterlund to rule out the possibility that women might shy away from competition because by winning the tournament they would impose a negative externality on the other group members.

Lastly, under Task 4, the participants were not presented with a new set of math problems. Rather, if this task was randomly selected for payment at the end of the experiment, a participant's compensation would depend upon the number of correct answers s/he provided under Task 1's scheme. In Task 4, a participant could choose which compensation scheme s/he wanted applied to his or her past performance in Task 1, piece rate or tournament. A participant would therefore effectively win a Task 4 tournament if his or her Task 1 performance had been the highest among the other participants in the group for that task. Before making their choices in Task 4, participants were reminded of their respective Task 1 performances. Thus, Task 4 allowed Niederle and Vesterlund to see whether gender differences in the choice of compensation scheme appeared even when no future and past tournament performance was involved. The authors could determine whether general factors such as overconfidence, risk, and feedback aversion caused a gender gap in the choice between the noncompetitive piece-rate and competitive tournament schemes. At the end of the experiment, before learning about their performance relative to the group's other participants, each participant was asked to guess her ranking (in terms of the number of correctly solved problems) in Tasks 1 and 2, respectively. Each participant picked a rank between 1 and 4 and was paid $1 for each correct guess.

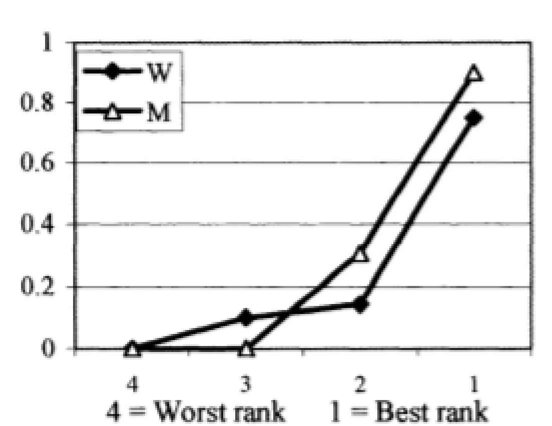
As expected, the authors found no statistically significant gender gap in performance under either the piece-rate or tournament schemes, with both sexes performing significantly better in a tournament. As a result, Niederle and Vesterlund conclude that there is no gender difference in the probability of winning the Task 2 tournament. More importantly for this particular experiment, the absence of a gender gap in the performance of Tasks 1 and 2 raises the expectation that a subsequent gender gap in Task 3 should likewise not be observed. However, this expectation was not fulfilled—35% of women and 73% of men selected the tournament, a statistically significant result. Moreover, as the figure below demonstrates, at each Task 2 performance level, men are more likely to enter the tournament in Task 3 (the figure's vertical axis measures the percentage of the group entering the tournament). Even women who score in the highest ("best") performance quartile chose to enter the tournament at a lower percentage than men in the lowest ("worst") performance quartile.



(Niederle and Vesterlund 2007)

In other words, while women shy away from competition, men are drawn to it.

Turning to Task 4, recall that although this choice is very similar to that of Task 3, Task 4's choice eliminates the prospect of having to subsequently participate in a competition. Thus, only in Task 3 could a gender gap in preference for competition have played a role in the choice of compensation scheme. As the figure below shows, there is no statistically significant gender gap in the choice of compensation scheme in Task 4 based upon perceived ranking in Task 1. A higher percentage of women than men who guessed their Task 1 ranking to be low (i.e., at level "3") chose the tournament scheme in Task 4, while the percentages were reversed for those participants who guessed their Task 1 rankings to be high (at levels "1" and "2"). But because the two lines in the figure remain close together, these differences are not statistically significant (i.e., we should treat the groups' respective choices as being no different from one another).



(Niederle and Vesterlund 2007)

This result from Task 4 cements the authors' finding that women shy away from actual competition slated to occur at a future point in time, not implicit competition based upon their interpretations of how their past performance compares with others.[10]

10. In a related study of the performances of men and women in professional judo fights for bronze medals (of all things!), Cohen-Zada et al. (2017) find that men's performances are significantly affected by what the authors' call "psychological momentum", while women's is not. Psychological momentum is defined as the tendency of an outcome (such as a win in an initial judo match) to be followed by a similar outcome (a win in a subsequent match) that is not caused by any strategic incentives of the players. The authors point out that this result is consistent with evidence in the biological literature that

TESTING FOR THE EXISTENCE OF AN ENDOWMENT EFFECT

Here is an experiment you and your fellow students can use to test the extent to which you exhibit an Endowment Effect—an experiment facilitated by your instructor, of course:[11]

> Ten of you have been given a coffee cup with your university's logo. You are henceforth known as "sellers." Ten of you have not been given this item. You are henceforth known as "Buyers." Each of the Sellers will now write on their piece of paper their "Seller's Price," the price at which they would willingly sell their coffee cup to one of the Buyers. Each of the Buyers will now write on their piece of paper their "Buyer's Price," the price at which they would willingly buy a coffee cup from one of the Sellers. Sellers' prices will be ranked from highest to lowest and then compared with the highest-to-lowest ranking of Buyers' prices to determine which trades will occur via paired bids.

For example, suppose from a class with 10 students, we randomly select five to be sellers and five to be buyers of the coffee cup. Each buyer and seller writes his or her price on a sheet of paper, and the outcome is tallied.

Seller/ Buyer #	Sellers Price	Buyers Price
1	2.00	2.20
2	1.50	1.50
3	1.30	1.20
4	1.00	0.50
5	0.50	0.30

In this "market," an equilibrium occurs where Buyer 1 pays Seller 1 \$2.00 for Seller 1's cup, Buyer 2 pays Seller 2 \$1.50 for Seller 2's cup, Buyer 3 pays Seller 4 \$1.00 for Seller 4's cup, and Buyer 4 pays Seller 5 \$0.50 for Seller 5's cup. Seller 3 does not end up selling and Buyer 5 does not end up purchasing a coffee cup. To see this, first note that Buyer 1's offer price (or willingness to pay (WTP)) exceeds Seller 1's asking price (or willingness to accept (WTA)), which is the largest WTA value among the group of sellers. Thus, Buyer 1 and Seller 1 are "matched," and Seller 1 is paid his WTA. Next, since Buyer 2's WTP just matches Seller 2's WTA (which is the next highest WTA value), Buyer 2 and Seller 2 are matched. The next highest WTP value is exhibited by Buyer 3, and since this value exceeds Seller

testosterone, which is known to enhance performance of both genders, typically increases following victory and decreases following loss only among men.

11. The inspiration for this experiment is drawn from Kahneman et al. (1990).

4's WTA (not Seller 3's), Buyer 3 is matched with Seller 4. Finally, Buyer 4's WTP just matches Seller 5's WTA, so Buyer 4 and Seller 5 are matched. In the end, no such matches can be found for Buyer 5 and Seller 3.

This is good information to have. But a question remains: is there evidence of an Endowment Effect in this market?

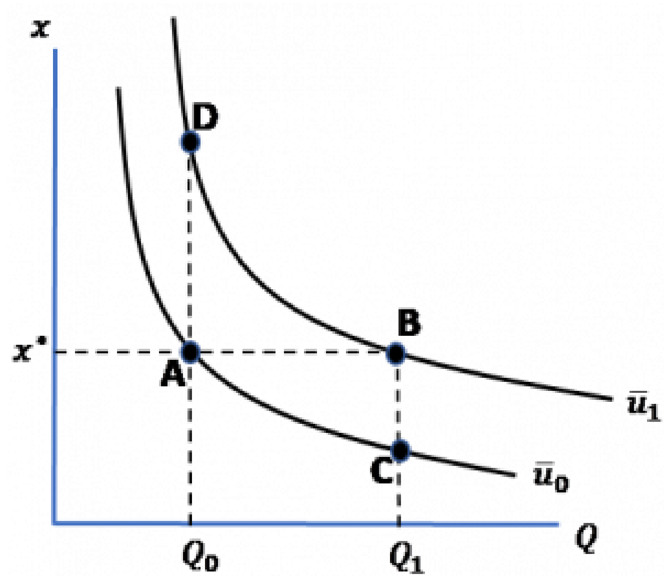
The answer is (most likely) "no" since four out of five possible sales were ultimately consummated. Although there is no hard-and-fast threshold for determining whether the effect has occurred, it seems safe to say that wherever one might put the threshold, four-out-of-five (or 80% of possible sales consummated) would lie above it. Or, to put it another way, in a market characterized by a relatively strong Endowment Effect exhibited by its sellers, one would expect most sellers' WTA values to exceed buyers' WTP values, resulting in few sales ultimately being made. Indeed, if pressed to nail down a threshold, one can really do no better than to metaphorically flip a coin (i.e., to choose a 50% threshold), which, in the case of our example market, means that only one or two consummated sales would have indicated the existence of an Endowment Effect.

HOMO ECONOMICUS AND THE ENDOWMENT EFFECT**

In Chapter 3 we introduced the graphical concept of an indifference curve. We mentioned that the stylized version of this curve—smooth, everywhere downward-sloping, and convex to the origin—can be used to represent *Homo economicus*' preferences over any two commodities. It turns out that this framework, as demonstrated in Hanley et al. (2007), can be used to reach an interesting conclusion about *Homo economicus*' susceptibility to the Endowment Effect.

In Figure 6.2 below, two indifference curves are drawn for our individual whom we'll name Ted, each curve corresponding to a different level of utility, $\bar{u}_0\$$ and $\bar{u}_1\$$. Recall that the relative locations of the two curves indicate that $\bar{u}_1 > \bar{u}_0$ (i.e., the curve drawn for utility level $\bar{u}_1\$$ corresponds to a set of bundles yielding a higher level of utility than the set of bundles denoted by the curve drawn for $\bar{u}_0\$$).[12]

Figure 6.2. *Homo economicus* and the Endowment Effect



12. Also recall that (1) because Homo economicus abides by the two rationality axioms introduced earlier (in particular, transitivity), and (2) if we further assume that Homo economicus' preferences are what's known as monotone, then the two indifference curves can never cross. The monotone property also ensures that an indifference curve never slopes upward.

For sake of example (but without loss of generality), Ted's two indifference curves have been drawn for two commodities denoted as x and Q. Commodity x is a "private good" that Ted has to purchase directly with his income. Commodity Q is a "public good" that Ted receives from the government without having to make a direct payment out of his income (e.g., more wilderness area for Ted to explore near his home, cleaner air for Ted to breath, etc.). We assume that commodities x and Q are the only two commodities Ted consumes.[13] Also identified in Figure 6.2 are points A – D, level x^* for commodity x, and levels Q_0 and Q_1 for commodity Q. Finally, for future reference we note that because Ted can only spend his income on commodity x, he ends up spending all of his income on, which means he's stuck with consuming level x^* regardless of whether he consumes level Q_0 or Q_1 of commodity Q. Therefore, the higher x^* is located on the vertical axis, the larger is Ted's income level, all else equal.

Okay. We are now ready to show why Ted, our *Homo economicus*, exhibits an Endowment Effect. We start by assuming that Ted initially consumes at point A (i.e., bundle (x^*, Q_0)) where he attains utility level $\bar{u}_0$. Now, let the level of the public good increase from Q_0 to Q_1. Note that because Ted is constrained to consume level x^* of commodity x, we know that he now consumes at point B where he attains the higher level of utility $\bar{u}_1$. It makes sense that Ted is happier at point B because he is now consuming the same amount of commodity x at x^* that he was originally consuming at point A, and he also gets to consume more of commodity Q (lucky him).

It turns out that vertical distance AC in Figure 6.2 represents Ted's WTP for this move from point A to point B. How so? Distance AC represents the maximum amount of commodity x that we could take away from Ted such that (1) his consumption of commodity Q is maintained at level Q_1, and (2) Ted is not left with fewer utils than his initial utility level, $\bar{u}_0$. Hence, vertical distance AC indeed represents Ted's WTP for the change in the level of good Q represented by $Q_1 - Q_0$ (measured in terms of commodity x).

To identify his WTA, we instead start by assuming that Ted initially consumes at point B (i.e., bundle (x^*, Q_1)) where he has attained utility level $\bar{u}_1$. Now, let the level of the public good decrease from Q_1 to Q_0 (i.e., some of the public good has been taken away from Ted—the wilderness area near his home has shrunk, or air quality has worsened, etc.). Note that because Ted is again constrained to consume level x^* of commodity x (because his income level has remained the same), we know that he now chooses to consume at point A where he regresses to the lower level of utility $\bar{u}_0$. It makes sense that Ted is less happy at point A because he is now consuming the same amount of commodity x at A that he was originally consuming at B and is, unfortunately, consuming less of commodity Q (woe to him).

Adopting a similar logic, vertical distance AD in Figure 6.2 represents Ted's WTA for this move from point B to point A. In this case, distance AD represents the minimum amount of commodity x we must give Ted such that this amount (1) holds his consumption of commodity Q at level Q_0 and (2) does not leave Ted with less than his initial utility level, $\bar{u}_1$. Hence, vertical distance AD indeed represents Ted's WTA for the change in the level of good Q represented by $Q_1 - Q_0$ (again, measured in terms of commodity x).

Recall that WTP and WTA both measure Ted's valuation of a given change in the amount of public good Q. However, the contexts within which the measurements occur are different. WTP presumes Ted does not initially own rights to the change in the amount of the good—it is his maximum

13. The facts that we have constrained Ted to consume only two commodities and that he is essentially receiving commodity Q for free, are innocuous assumptions made for the sake of our simple example.

willingness to pay to gain ownership of that changed amount. In contrast, WTA does presume that Ted initially owns rights to the change in the amount of the good—it is his minimum willingness to accept the loss of ownership of that changed amount. Thus, if Ted's WTA exceeds his WTP for the same amount of change in good Q, then he exhibits an Endowment Effect. This is because he would then place a higher value on amount Q_1 relative to Q_0 when he initially owns Q_1 as opposed to when he does not. In other words, Ted places a higher value on the change in the amount of Q simply because he is endowed with the right to own that changed amount. Clearly, distance AD exceeds distance AC in Figure 6.2, implying Ted's WTA is greater than his WTP for the given change in the amount of good Q. Thus, even a *Homo economicus* like Ted exhibits an Endowment Effect.[14]

CONCLUDING REMARKS

Recall that the goal of Section 2 is to explore exactly where the standard, rational-choice theory ascribed to *Homo economicus* fails to adequately predict behavior we repeatedly see exhibited by *Homo sapiens* in laboratory and field experiments. We have learned that, in several key respects, this behavior systematically violates the theory. The violations appear in a myriad of forms; as violations of axioms known as Dominance, Invariance, Independence, and Substitution, and the Sure-Thing Principle. From these violations spring two questions—why and toward what ends? The question "why" pulls us back into the realms of cognitive and psychological sciences, realms rich in reminders of what makes *Homo sapiens* so exquisitely quirky and idiosyncratic and inherently fallible. And the question "toward what ends" propels us forward into the realm of understanding where and how rational-choice theory fails to adequately explain human choice behavior. In this realm, the new theories of behavioral economics arise.

Following Kahneman (2011) and others, we call these quirks and idiosyncrasies, "miscalculations, biases, fallacies, heuristics, and effects." We recall specific terminology such as Affect and Availability Heuristics, Priming and Framing Effects, Status Quo and Confirmation Biases, Conjunction Fallacy, the Law of Small Numbers, conformity, mental accounting, and the overweighting of improbable events, among others. These are the types of human foibles distinguishing *Homo sapiens* from *Homo economicus*. They materialize as violations of rational choice theory. As we have learned, the violations were identified in the foundational laboratory experiments conducted by Nobel Laureates Daniel Kahneman and Richard Thaler, among many others, experiments that you and your classmates have now participated in yourselves.

It is one thing to identify where a body of theory fails to adequately explain real-world phenomena. It is another to revise that theory in an effort to advance not only the theory but also to gain a deeper understanding of the human experience itself. Toward this end, we have learned about the value function and how it can be used to depict such behaviors among *Homo sapiens* as reference dependence, loss aversion, and the Endowment Effect. And we have learned about the complications imposed on the rational-choice theory by human emotions such as envy, guilt, regret, and blame, as well as *Homo sapiens*' discomfiture with ambiguous circumstances and the need to feel and appear competent.

14. Interestingly, to the extent that Ted's indifference curves are linear (i.e., are still downward-sloping, but with constant slopes rather than the non-constant slopes depicted for the stylistic case in Figure 6.2), his WTP and WTA values (i.e., distances AC and AD in the figure) converge toward equality.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013).

1. † Which of the two utility functions represents transactional utility and why? $u\left(x\right) = 5x - x^2 - k\left(x\right)$ and $u\left(x\right) = 5x + \frac{x}{k(x)} - x^2 - k\left(x\right)$, where x represents the total amount of a good purchased, and $k(x)$ represents the total cost incurred by the consumer in purchasing the good.

2. Why might people be more likely to spend a small inheritance and invest a large one? Which quirk of *Homo sapiens* introduced in this chapter does this likelihood represent?

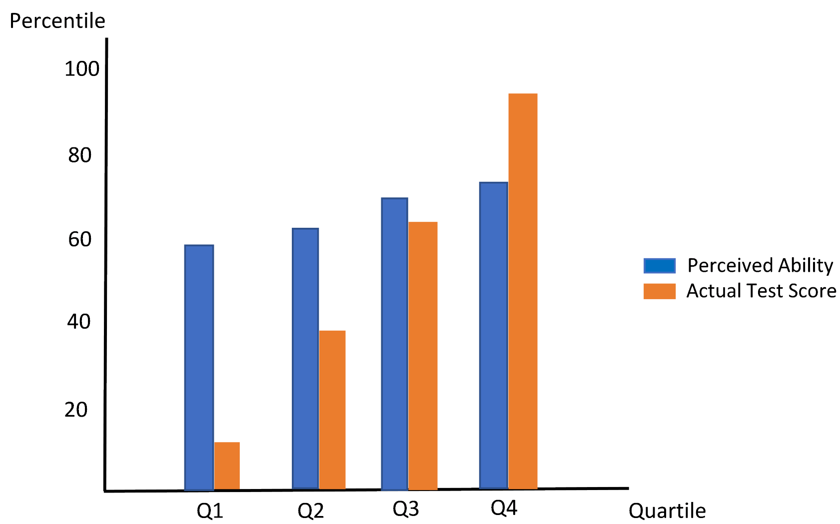
3. † How does the use of a credit card confound the conclusion reached by Prelec and Loewenstein (1998); namely that, as an example of mental accounting, *Homo sapiens* seem to prefer decoupling payments for durable goods but not necessarily for non-durable (perishable) goods?

4. Can you think of an example from your own life where you have practiced mental accounting? Explain.

5. † Ella is a high school student who is given $3 each school day by her parents to be used as "lunch money." She works a part-time job after school, earning a small amount of "spending cash." In addition to her lunch money, Ella spends $5 from her spending cash each week on lunch. Suppose her parents reduced Ella's lunch money by $2 per day but that she simultaneously receives a $10-per-week raise at her job, requiring no extra effort on her part. What would the rational choice model suggest should happen to Ella's spending on lunch? Alternatively, what does the mental accounting framework predict?

6. Each year, hundreds of thousands of people receive coronary-artery bypass surgery. The surgery saves their lives, but only for the long term if they adopt some important lifestyle changes such as dietary changes, quitting smoking, more physical exercise, and managing stress more effectively. Suppose Hayden the Hyperbolic Discounter, who has a hankering for eating junk food, smokes cigarettes occasionally when stressed out, and avoids regular exercise, decides to have coronary bypass surgery. What do you think the long-run outcome is going to be for Hayden?

7. Which experiment discussed in this chapter do these results pertain to?



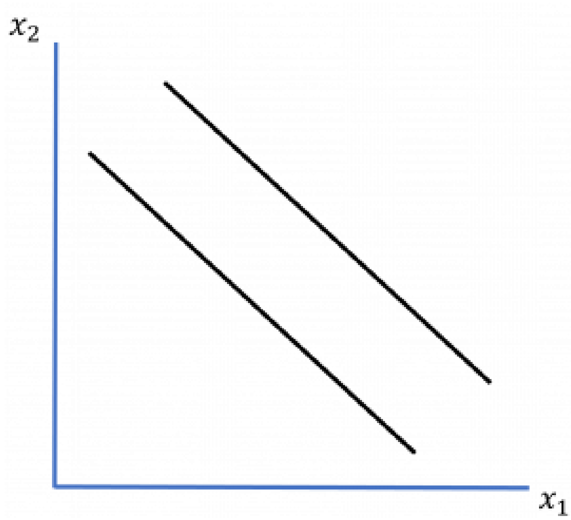
8. Suppose Evelyn the Environmental Economist is presenting her case in a public meeting for why raising the price of municipal water in the face of persistent drought conditions would be a good thing for the community, when someone in the audience yells out, "That's unfair for seniors and others living on fixed incomes." How might Evelyn frame her response in a way that dispels the audience's concerns about the fairness of a price increase?

9. How would the indifference curve in Figure 6.1 change when drawn for a person who suffers from guilt but not envy? Draw the curve.

10. Can you recall an example from your own life where you exhibited an Endowment Effect that ultimately led to regret?

11. The Gender Gap experiment discussed in this chapter measured gender differences in terms of how males and females deal with competitive situations. Think of another situation where a gender gap may exist and design an experiment to test for it.

12. It was shown in this chapter that a *Homo economicus* who exhibits convex-shaped indifference curves exhibits an Endowment Effect. Does this result still hold if *Homo economicus* exhibits linearly shaped indifference curves, as depicted in the figure below? Show your result using this graph.



13. Explain the relationship between the Decoy Effect and reference dependence. Describe a situation in your own life where you have been a victim of the Decoy Effect.

14. Describe a negative implication of the Zero-Price Effect in the real world. Have you ever fallen victim to this effect in your own life? If so, describe the circumstances.

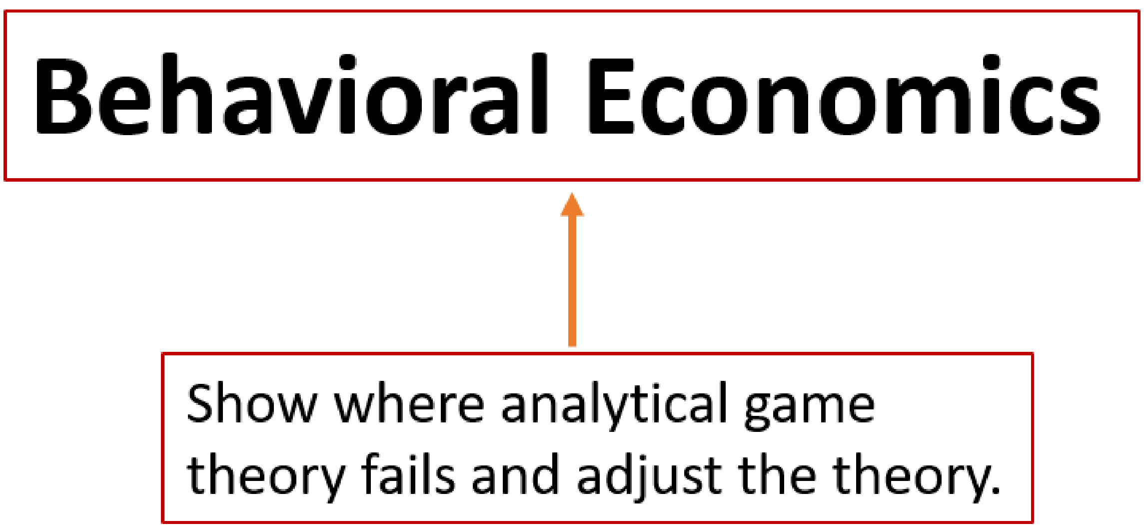
Media Attributions

PART III.

SECTION 3 - BEHAVIORAL GAME THEORY

Sections 1 and 2 covered a big chunk of territory in the field of behavioral economics, where the focus was on identifying and explaining individual behavior in general. We now venture into the equally enriching territory of behavioral game theory where the focus is again on identifying and explaining human behavior. This time, however, we consider individual behavior in a social setting and, from time to time, assess the advantages and disadvantages of teams of individuals.

Referring to the diagram presented in the This Book's Approach section, Section 3 of the textbook pertains to the diagram's bottom portion.



Here, we again demonstrate how standard economic theory (in this case, game theory) fails by highlighting the major disconnects between the behavior predicted of *Homo economicus* and that actually displayed by *Homo sapiens* in social settings. By "social settings," we mean game-like situations where two or more individuals compete against each other for payoffs of some kind (typically monetary payoffs, but not always). Most of the games share a tantalizing aspect—analytically speaking, the games are structured such that their equilibria correspond to the individuals having chosen not to cooperate with each other. These are the equilibria we expect *Homo economicus* to reach. Cooperation, if it occurs, is typically evinced by *Homo sapiens*, not *Homo economicus*. Are you surprised?

One thing you will notice as you learn about the games and participate in them yourselves, is that while it does as equally an effective job as the experimental economics of Kahneman, Tversky, Thaler, et al. in revealing the extent to which *Homo sapiens* violate the tenets of rational choice, the subfield of behavioral game theory tends not to propose new or revised theories *per se*. Rather, the main contributions of behavioral game theory are found in the innovative ways in which the original, foundational games have been tweaked over time to account for those violations.

Unless otherwise indicated, the games presented in this section are based on discussions presented in Camerer (2003).

Media Attributions

CHAPTER 7.

SOME CLASSIC GAMES OF ITERATED DOMINANCE

Before diving into the deep pool of behavioral game theory, we need some specific nomenclature about what constitutes a game and its solution, or what we have been calling its equilibrium. If you've ever played a board or card game with your friends or family, then none of this terminology should surprise you.

A game consists of a set of "players," each with their own set of "strategies." Precise "rules" govern the "order" in which players make their "moves," the "information" they have available, and, ultimately, their "payoffs." I don't know about you, but the card game poker comes immediately to mind. The keywords are players, strategies, rules, order, moves, information, and payoffs.

We expect that *Homo economicus* will attain what's known as a Nash equilibrium, or perhaps a refinement of Nash equilibrium, depending upon the game being played.[1] Simply put, a Nash equilibrium prevails when each player can no longer adjust his or her strategy to obtain added payoff. Thus, in a Nash equilibrium, all players have chosen respective strategies that are the best responses to each of the other players' strategies. The Nash equilibrium is derived analytically and, thus, is highly predictable. We will see just how predictable the equilibrium is in a wide variety of games. Since this is the equilibrium obtained by *Homo economicus*, we henceforth use the terminology *Homo economicus* and "analytical equilibrium" inter-changeably.

As we will learn, the equilibria typically obtained in games played by *Homo sapiens* expand upon the Nash equilibrium concept by adding in such aspects of the human experience as emotion, miscalculation, limited foresight, doubt about how informed the other players are, and learning-by-doing—many of the same human quirks and idiosyncrasies we encountered in Section 1. The equilibria obtained in games played by *Homo sapiens* are typically derived more intuitively than analytically. Thus, the equilibria are generally unpredictable.

Let's start with one of the most famous and basic of games—ultimatum bargaining.

ULTIMATUM BARGAINING

Consider the following game presented in Camerer (2003):[2]

1. The Nash equilibrium solution concept is attributed to John Nash Jr., an American mathematician and the winner of the 1994 Nobel Memorial Prize in Economic Sciences. Nash published his pioneering work in non-cooperative game theory in the early 1950s (Nash, 1950a, 1950b, 1951). His struggles with mental illness and recovery are recounted in Sylvia Nasar's 1998 biography A Beautiful Mind and later, the 2001 film of the same name.

2. Just (2014, pages 420-422) provides a mathematical framework within which to assess the same outcomes of this game as we demonstrate here with more intuitive game-theoretic reasoning.

Two players—a Proposer and a Responder— bargain over \$100. The Proposer offers some portion, x, of the \$100 to the Responder, leaving the Proposer with \$(100 – x). If the Responder accepts the offer, then she gets \$$x$ and the Proposer gets \$(100 – x). If the Responder rejects the offer, both players get nothing.

The analytical equilibrium for this game evolves according to the following logic:

> *By going first, the Proposer possesses all of the bargaining power. The Proposer, therefore, exploits the fact that the self-interested Responder will take whatever is offered. The amount offered by the Proposer is thus very close to zero. Surmising that this is indeed the Proposer's best strategy, and also recognizing that he gets nothing if he rejects the Proposer's offer, the Responder has no better strategy than to accept whatever the Responder offers, as meager as the offer is. The Proposer knows that this is the logic the Responder will use, and the Proposer knows that the Responder knows this, and so on. Hence, the analytical equilibrium is that the Proposer makes the meager offer (in the limit, \$0.01) and the Responder accepts.*

Ouch. Before exploring what the behavioral game theory literature has to say about how *Homo sapiens* have actually played this game (i.e., what equilibria they have obtained), it is informative to link this game to the nomenclature presented at the chapter's outset.

The players are a Proposer and Responder. The Proposer's strategy is to choose an offer amount, x, that he thinks will ultimately be accepted and result in a desired payoff amount. The Responder's strategy is to accept or reject the Proposer's offer. The rules of the game, which govern which player moves when (i.e., the order of moves) and how the resulting payoffs are determined, are clearly spelled out. The Proposer moves first by making offer x and the Responder moves second, choosing to accept or reject the offer. After the Responder's decision is made, the payoffs are distributed according to the following rule: If the Responder accepts the Proposer's offer, the payoffs are x for the Responder and $(100 - x)$ for the Proposer; if the Responder rejects the Proposer's offer then the payoffs are zero for each.

As the logic behind the determination of the analytical equilibrium makes clear, the information available to the Proposer and Responder has an important bearing on the game's analytical equilibrium. Although this game has allocated all of the bargaining power to the Proposer, both the Proposer and the Responder are assumed to share complete and common information. Each player not only knows what payoffs he stands to gain via the game's rule, but also what payoffs the other player stands to gain, and each player knows that the other player knows this, and so on. The fact that all players know the same things about the game is what is common about the information. The fact that no information is hidden from the players is what makes the information complete.

The solution process for the Ultimatum Bargaining game's analytical equilibrium follows what's known as "iterated dominance" due to (1) the players making their moves sequentially (or iteratively), and (2) the concomitant need for each player to think ahead about the other player's subsequent move before choosing what to do presently. In this case, because each player's best strategy is calculatable and unique, we say that it is dominant.[3] Further, because the Proposer in this game initially considers

3. An extreme form of Ultimatum Bargaining is known as the Dictator Game, whereby the Proposer makes an offer that the Responder must accept (i.e., the Responder is not allowed to reject). There is no iteration and no real role or advantage for

what should happen in the last stage, where the Responder decides whether to accept or reject the offer made in the first stage, iterated dominance is operationalized via "backward induction." The Proposer first figures out what should be the outcome of the game's final stage and then works back from there to determine what she should do in each preceding stage all the way back to the first stage. Because the equilibrium is solvable via backward induction, we say that it is "subgame perfect."[4] We will be seeing examples of backward induction and subgame perfection repeatedly in this chapter, so get ready!

From an analytical, game-theoretic perspective, this is all interesting to know. But what about *Homo sapiens*? How have we actually played the Ultimatum Bargaining game? We have several different kinds of results. Camerer (2003) has compiled an exhaustive list of studies that have considered ultimatum bargaining with varying rules, payoff amounts, and multiple rounds, in different regions of the world with varied cultural contexts, with men vs. women, and more. He concludes that results from the different versions of the game are quite robust. Modal and median ultimatum offers are usually 40%–50% of the total amount available to bargain over, and means are 30%–40%. There are hardly any offers made by the Proposer in the outlying category of 0%-10%, and the hyper-fair category 51%–100%. Offers of 40%–50% are rarely rejected. Offers below 20% or so are rejected about half the time (Camerer, 2003).

In other words, *Homo sapiens* do not generally converge to the game's analytical equilibrium. It seems that Proposers are susceptible to emotions like guilt, fairness, and/or altruism, and Responders succumb to envy and fairness (in this case, "reciprocity"). Here is a taste of some of the findings:

- Ironically, participants in more-primitive cultures in Africa, the Amazon, Papua New Guinea, Indonesia, and Mongolia have been found to behave more like *Homo economicus* than do participants in less-primitive cultures in the US, Europe, and Asia (c.f., Slonim and Roth, 1998; Buchan et al., 2004; Henrich et al., 2001 and 2002; Henrich, 2000).
- Repeated games with "stranger matching" and no provision of "history of moves" show a slight tendency for both offers and rejections to fall over time. Provision of history correlates with more pronounced reductions in offers and rejections (c.f., Roth et al., 1991; Bolton and Zwick, 1995; Knez and Camerer, 1995; Slonim and Roth, 1998; List and Cherry, 2000).
- Responders are not necessarily more likely to reject, say, \$5 out of \$50 than \$5 out of \$10, and similarly 10% of \$50 than 10% of \$10. In other words, the game's stakes do not necessarily matter (c.f., Camerer and Hogarth, 1999; Roth et al., 1991; Forsythe et al., 1994; Hoffman et al., 1996; Straub and Murnighan, 1995; Cameron, 1999; Slonim and Roth, 1998).
- Male Proposers do not necessarily offer more to attractive female Responders, but female Proposers have been found to offer more to attractive male Responders (Hamermesh and Biddle, 1994).

having information. Yet, there is a dominant strategy for the Proposer, which results in an analytical equilibrium where the Proposer offers nothing to the Responder (i.e., $x = 0$) and the Responder accepts. Obviously, in this game, "Responder" is a euphemism for "Lackey." See Forsyth et al. (1994), Hoffman et al. (1994), Hoffman et al. (1996), and Fehr and Fischbacher (2004) for experiments with the Dictator Game.

4. Technically speaking, a subgame perfect equilibrium is obtained when a Nash equilibrium has been reached at every subgame of the original game, even if a particular subgame has not been played. In the case of Ultimatum Bargaining, there are two subgames: one where the Responder either accepts or rejects the offer, and the other is the full game itself (the full game is always considered a subgame). We will learn more about what defines a "subgame" a bit later in the chapter.

- Young children are more self-interested, *Homo economicus*-like Proposers and Responders, but then become more fair-minded as they grow older (Damon, 1980; Murnighan and Saxon, 1998; Harbaugh et al., 2000).
- Calling the game a "seller-buyer exchange" encourages self-interest. Describing the game as a "common pool resource" encourages generosity (Hoffman et al., 1994; Larrick and Blount, 1997). Note that this is an example of a framing effect!
- When Proposers know the exact amount of money to be divided, and Responders either know nothing at all or know the probability distribution of possible amounts, Proposers offer less (c.f., Huck, 1999; Camerer and Loewenstein, 1993; Mitzkewitz and Nagel, 1993; Straub and Murnighan, 1995; Croson, 1996; Rapoport et al., 1996). This is a consequence of "incomplete information." However, when Responders know the alternative amounts that the Proposer could have offered, they tend to exhibit "inequality aversion" (or, alternatively, a commitment to fairness) and reject the Proposer's offer (Falk et al., 2003).
- Creating a sense of entitlement by letting the winner of a contest (played beforehand) be the Proposer lowers offers (c.f., Hoffman et al., 1994; List and Cherry, 2000). This is known as an Entitlement Effect.

Raworth (2017) eloquently sums up the main takeaway from these disparate findings: *Homo sapiens*' sense of reciprocity appears to co-evolve with their economy's structure, or if you like, the context within which the game is played. In addition to the varied contexts described above, the structure of the Ultimatum Bargaining game has been modified as well. We consider two of these structurally adjusted versions of the game—the Nash Demand Game and the Finite Alternating-Offer Game.

NASH DEMAND GAME

Consider the following game proposed by Mehta, et al. (1992):

> This game has three stages. Ultimately, at the third and final stage, the two players individually state their "demands." If the two demands add to \$10 or less, then they each get their individual demands; otherwise, they each get nothing. In the first stage, the players are each dealt four cards randomly from a deck with eight cards only: four aces and four deuces (i.e., twos). Players are told that if all four aces are ultimately held by one player, then that player's cards are worth \$10, in which case each player will have earned the right to state his demand in the third and final stage. Otherwise, with any other configuration of aces held by the two players, each player's cards are worth nothing. In the second stage, the players trade their cards with each other.

The analytical equilibrium for this game obtains evolves according to the following logic:

> *Since Homo economicus know the composition of the deck, one player can tell from his own hand how many aces the other player has—namely, four minus his own number of aces. Thus, in the second stage, the players should always trade with each other such that the four aces end up being held by one of the players, as this gives them the right to state their demands in the third and final stage. Recognizing that*

the cards were randomly dealt to begin with, the players should each state a demand of $5 in the final stage.

What happens when *Homo sapiens* play the Nash Demand (ND) game instead? Mehta, et al. provide an answer. The authors start by considering what happened when Player 1 was dealt two aces in the first stage of their experiment. Of the 42 instances where this happened, 40 resulted in the player ultimately demanding half the pie of $10 in the final stage of the game. Thus, Player 1 behaved as would be expected of *Homo economicus*. However, when Player 1 was dealt either one ace or three aces, she demanded half the pie only roughly half of the time—16 of 32 times when dealt one ace and 17 of 33 times when dealt three. The other half of the time, Player 1 demanded a fraction roughly equal to the fraction of aces originally held—16 of 32 times when dealt one ace and 15 of 33 times when dealt three. As a result, there is a 22% ((16/32) x (15/33) = 0.22) deviation from the analytical equilibrium (of an even split in demands) in cases where one and three aces have been dealt to Player 1.

Mehta, et al. postulate that the implicit information about how many aces each player originally held in the ND Game created "focal points" for this type of deviation. For example, when Player 1 was dealt one ace and three deuces she was able to discern that Player 2 held the other three aces and one deuce. In those instances, Player 1 determined that because she contributed only one of four aces now held by Player 2 she should demand less than half of the $10. Similarly, when Player 1 was dealt three aces she should demand more than half of the $10.

In a slight twist on the basic ND game, Binmore et al. (1998) had their subjects play the game with an "outside option." This game is played according to the same rules as the basic ND game except that before the game begins, Player 2 is randomly given a commonly known outside option worth $0.90, $2.50, $4.90, $6.40, or $8.10. In other words, before the game begins, Player 2's outside option (which is equal to one of the five possible values stated in the previous sentence) is announced to both players. Player 2 can choose to take the option, in which case he gets that payment, and Player 1 gets nothing. Or Player 2 can turn down the option, and the ND game ensues.

A *Homo economicus* version of Player 1 should ignore Player 2's outside option since if Player 2 turns down the option and opts to play the ND Game, then playing the game from that point forward is all that matters. Being a member of *Homo economicus*, Player 2 knows that this is indeed Player 1's best strategy, and thus, if the ND game is ultimately played, Player 1 will demand $5, which means that the most Player 2 will be able to demand is also $5—the analytical equilibrium is therefore obtained. Hence, Player 2 will take the outside option only if it is worth $6.40 or $8.10. Otherwise, Player 2 should turn down the option and play the ND game with Player 1.

Binmore et al. found that Player 2s do not behave like *Homo economicus*. For instance, one-third of Player 2s opt out at the option value $4.90 and only 60% opt out at $6.40. Further, the demands of those Player 2s who opt-in at option values above $5.00 match those (focal) values rather than the expected $5.00. Player 1 behaviors deviate less from what we would expect of *Homo economicus*. Their demands are relatively close to $5.00 except in cases where Player 2's option values exceed $5.00. Interestingly, Player 1's demands decrease in accordance with the commonly known option values for Player 2, resulting in a total demand less than the $10 threshold. To the extent that Player 1 expects Player 2 demands to tilt toward the focal points of their option values, Player 1 is actually making a rational choice in lowering her demands. And to the extent that Player 2 with higher option values expects Player 1 to lower her demand accordingly, then Player 2 is likewise making a rational choice. The fact that relatively few Player 1s make demands that leave Player 2s with less than their option values is also rational to the extent that Player 1s expect Player 2 demands to tilt to their option values.

So, while the behavior of the players in Binmore et al.'s (1998) ND game with an outside option does not adhere to those expected in an analytical equilibrium, to the extent that their behaviors are premised on *Homo sapiens*' tendencies to tilt toward focal points in these types of games we can interpret the players as nevertheless making contextually rational choices.

FINITE ALTERNATING-OFFER GAME

Consider the following game presented in Camerer (2003):

> Two players bargain for two periods. In the first period, Player 1 offers a division of $200 to Player 2. If Player 2 rejects Player 1's offer, the "pie" of $200 shrinks to $50 and Player 2 makes a counteroffer to Player 1. If Player 1 rejects Player 2's counteroffer, the game is over and neither player gets anything.

Solving this game analytically requires the use of backward induction, which results in a subgame perfect equilibrium (SPE). The logic goes like this:

> *Using backward induction, Player 1 considers what Player 2 will do in the second period when it is Player 2's turn to make the counteroffer. Player 1 does not want the initial offer to be rejected since this will shrink the pie to $50. So, Player 1 offers at most $50 to Player 2 ($50 being the most Player 2 could ever hope to get if he rejects Player 1's offer). Player 1, therefore, keeps at least $150 and Player 2 gets at most $50.*

Camerer reports that in games played with *Homo sapiens,* Player 1 tends to offer half of the pie in the first stage (e.g., out of a sense of fairness or fear that the initial offer might otherwise be rejected by Player 2), in which case Player 2's ability to reject the initial offer is perceived as a credible threat by Player 1. However, with repeated play, Player 1 quickly learns to offer the SPE amount of $50 in the first stage. In other words, as *Homo sapiens* learn how to play the game, Player 2's threat is no longer perceived as being all that credible. Incredible, huh? With learning, *Homo sapiens* attain the analytical equilibrium.

CONTINENTAL DIVIDE GAME

Consider the following game proposed by Van Huyck et al. (1997):

> Players (more than two) each pick a number from 1 to 14. The rows of the matrix below show each player's payoff (in dollars) corresponding to the number she has chosen and the median choice made by the group as a whole.

	Median Choice													
Choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	45	49	52	55	56	55	46	-59	-88	-105	-117	-127	-135	-142
2	48	53	58	62	65	66	61	-27	-52	-67	-77	-86	-92	-98
3	48	54	60	66	70	74	72	1	-20	-32	-41	-48	-53	-58
4	43	51	58	65	71	77	80	26	8	-2	-9	-14	-19	-22
5	35	44	52	60	69	77	83	46	32	25	19	15	12	10
6	23	33	42	52	62	72	82	62	53	47	43	41	39	38
7	7	18	28	40	51	64	78	75	69	66	64	63	62	62
8	-13	-1	11	23	37	51	69	83	81	80	80	80	81	82
9	-37	-24	-11	3	18	35	57	88	89	91	92	94	96	98
10	-65	-51	-37	-21	-4	15	40	89	94	98	101	104	107	110
11	-97	-82	-66	-49	-31	-9	20	85	94	100	105	110	114	119
12	-133	-117	-100	-82	-61	-37	-5	78	91	99	106	112	118	123
13	-173	-156	-137	-118	-96	-69	-33	67	83	94	103	110	117	123
14	-217	-198	-179	-158	-134	-105	-65	52	72	85	95	104	112	120

(Van Huyck, et al. 1997)

For example, if a player chooses 4 and the median is 5, the player earns a healthy payoff of \$71. If the median is instead 12, the player earns -\$14 (i.e., she loses \$14)

Before discussing the logic behind the game's analytical equilibrium, it is useful to point out where backward induction and subgame perfection factor into determining the equilibrium, if at all. It turns out that backward induction is actually a moot point in this game. This is because there is only a single stage—all players simultaneously choose their numbers, which then automatically determines the median number and attendant payouts. One might be tempted to say that there are $14n$ subgames, where n represents the number of different players. This is not correct. There are instead $14n$ possible outcomes to what is only a single subgame (the game itself).

The logic behind the game's analytical equilibrium goes like this:

> *First, note that for any median less than or equal to 7, a player's best response is to choose the number 3. This is because, being a member of Homo economicus, each player knows that every other player is both self-interested and thinking the same way. Thus, if each player chooses 7—which results in a median of 7—it will be in a given player's self-interest to deviate and choose the number 5. But every player is equally self-interested and thinks the same way. Thus, a median of 5 results. But at a median of 5, each player deviates to the number 4. And at a median of 4, each player deviates to 3. Only at the choice of 3 does this madness stop. We call this the "low" Nash equilibrium. Using the same logic, for any median greater than 7 a player's best response is to choose the number 12. We call this the "high" Nash equilibrium. We expect this game's analytical equilibrium to be the high Nash equilibrium.*

Van Huyck et al. played this game with 10 different groups of *Homo sapiens*, each group playing the game 10 times in a row. What they found were basins of attraction. For groups that start with a median of 7 in the first period, the equilibrium converges to medians of 3, 4, and 6 (i.e., there is an attraction toward the low Nash equilibrium). For groups that begin with a median greater than 7 in the first period, the basin of attraction leads toward medians of 12 and 13 (i.e., the high Nash equilibrium). Hence, while not all groups of *Homo sapiens* obtain the high Nash equilibrium, it seems that roughly half do. The outcome is what we call "path dependent"—dependent upon where the path begins.[5]

BEAUTY CONTEST

Consider the following game presented in Camerer (2003):

> Each of N players simultaneously chooses a number x_i in the interval [0, 100]. The average for the group is calculated and then multiplied by a factor p <1 (say, p = 0.7). The player whose number is closest to this target (in this case, 70% of the average) wins $20.

Probably like you, the relationship between this game and a beauty contest escapes me.[6] But the game, regardless of what we call it, provides a nice example of how iterated dominance can be used to identify an analytical equilibrium.[7] The logic for the analytical equilibrium is as follows:

> *Each player starts by thinking, "Suppose the average is 50." Given this, he chooses the number 35 (0.7 x 50). But he would not stop here (i.e., he would begin iterating). He realizes that everyone else is making the same calculation, so he will choose 25 instead (0.7 x 35). But wait. He would then choose 18 (0.7 x 25). But wait……he would ultimately choose zero, which is this game's analytical equilibrium.*

Camerer reports results for one Beauty Contest played by groups of *Homo sapiens*, where p = 0.7 and there are low stakes of $7 and high stakes of $28. Irrespective of the stakes, *Homo sapiens* do not generally converge to the analytical equilibrium where each player chooses zero. But *Homo sapiens* do get close, especially when the stakes are higher. Camerer (2003) also reports that in most studies, players have used anywhere from zero to three levels of iterated dominance, which, according to the logic for the analytical equilibrium, means that the numbers most frequently chosen are 50, 35, and 25—quite a ways from zero.

TRAVELER'S DILEMMA

Consider the following game proposed by Capra et al. (1999):

> Two players simultaneously state price claims, between $300 and $750, for luggage lost by their airline company. The airline pays both players the minimum claim. The airline also adds a reward

5. Gladwell (2002) recounts a classic example of path dependence—The Broken Window Theory. The theory is that a single broken window left unfixed in a neighborhood can lead to a spiraling process of social breakdown as those with criminal intent interpret the broken window as a signal that the neighborhood is in decline and thus less able to protect itself. The broken window sets the neighborhood on the path toward a low Nash equilibrium.
6. Keynes (1936) described the action of rational agents in an equity market using the analogy of a fictional newspaper contest where participants were instructed to choose the six most attractive faces from among a hundred photographs. Those who chose the most popular faces were then eligible for a prize.
7. Note that this is a game where backward induction is again moot, and the only subgame is the game itself.

of $50 to the player who states the lower claim, and subtracts a penalty of $50 from the player who states the higher claim.

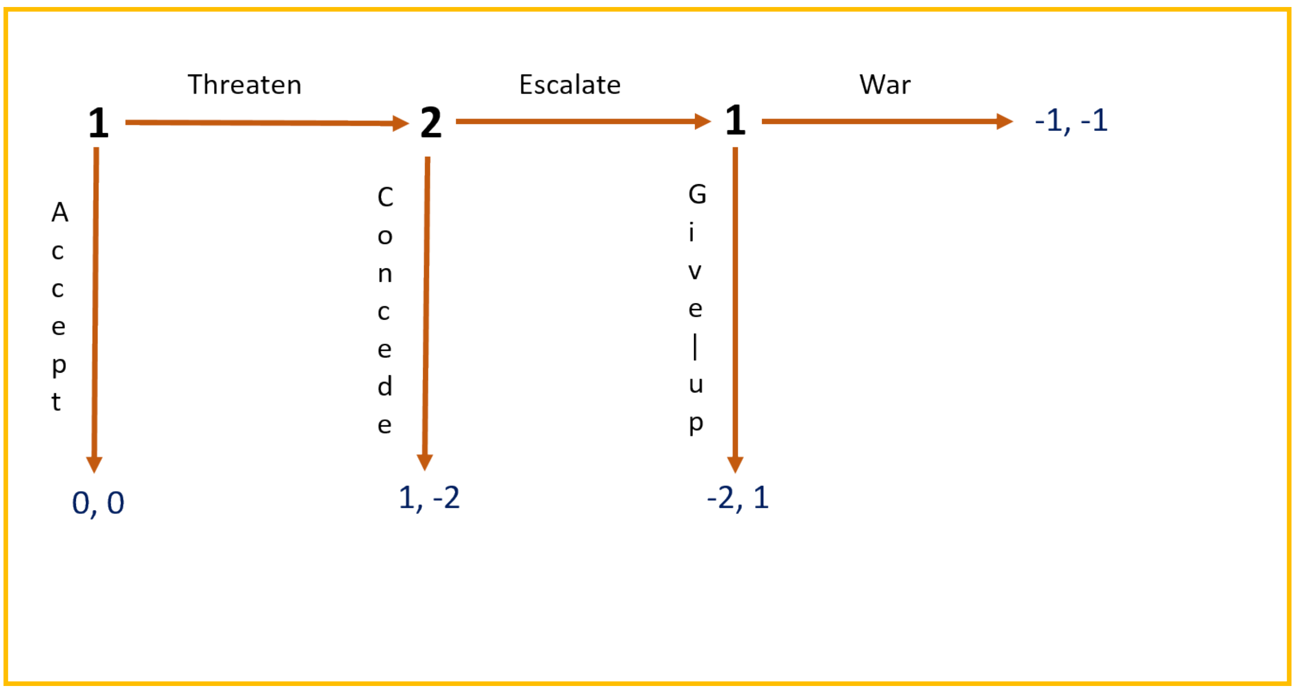
Applying iterated dominance, the logic for the analytical equilibrium goes like this:

> *Players should state claims that are one cent below what the other player is expected to state. In this way, a player helps boost the minimum claim (and hence her payoff) while earning the $50 reward. The result is a race to the bottom in which both players end up choosing the minimum claim of $300, and thus, neither player wins the reward (or, thankfully, suffers the penalty). This is the game's unique Nash equilibrium.*

Homo sapiens? Capra et al. found convergence toward the analytical equilibrium with their subjects over 10 periods of play only for the higher reward/penalty levels. In the later periods, average equilibrium claims were inversely related to the reward/penalty levels (the lower the reward/penalty level, the higher the average claim). Once again there is some evidence to suggest that *Homo sapiens* learn to converge toward (not necessarily all the way to) the analytical equilibrium, and the stakes of the game matter to some degree.

ESCALATION GAME

Spaniel (2011) explores the Escalation Game, depicted below as a decision tree.[8]



There are two players in this game—Player 1 and Player 2. In the first stage, Player 1 decides whether to "Threaten" Player 2 or "Accept." If Player 1 accepts, the game ends with both Players 1 and 2 receiving payouts of $0 each (the number to the left of the comma denotes Player 1's payout, and

8. Decision trees in game theory are known as games depicted in "extensive form." Drawing a decision tree is typically the most effective way to depict a multi-stage game.

the number to the right denotes Player 2's). If Player 1 chooses to threaten Player 2 in the first stage, then the game proceeds to the second stage where Player 2 gets to choose whether to "Escalate" or "Concede." If Player 2 concedes, the game ends with Player 1 receiving a payout of \$1, and Player 2 is required to make a payment to the experimenter of \$2. If Player 2 chooses to escalate in the second stage, then the game proceeds to the third and final stage where Player 1 gets to choose "War" or to "Give up." If Player 1 chooses Give up, then the game ends with Player 1 making a payment of \$2 to the experimenter and Player 1 receiving a payout of \$1. If Player 1 instead chooses War, then the game ends with both players required to pay the experimenter \$1 each.[9]

Can you guess the logic behind the game's analytical equilibrium?

> *Via backward induction, we start at the third and final stage and work our way back to the first stage. We see that Player 1 will declare "war" if the game ever reaches its final stage since paying \$1 is a better outcome for Player 1 than paying \$2. Knowing this, Player 2 will choose to "escalate" in the penultimate stage since she will be required to pay \$1 as a consequence of war occurring in the final stage, which is a better outcome for Player 2 than paying \$2. But then knowing this, Player 1 will choose to "accept" in the first period, which leads to a zero payout, which is, nevertheless, a better outcome than the payment of \$1 Player 1 would be required to pay as a result of later going to war with Player 2. This is the game's unique SPE.*

Note that in the case of international relations, this game captures the essence of "mutual deterrence." What drives mutual deterrence in the context of this game is that Player 2 choosing to escalate in the penultimate stage acts as a credible threat to Player 1.[10] What's the outcome when you and your classmates play this game? Hopefully, you choose mutual deterrence as opposed to going to war.

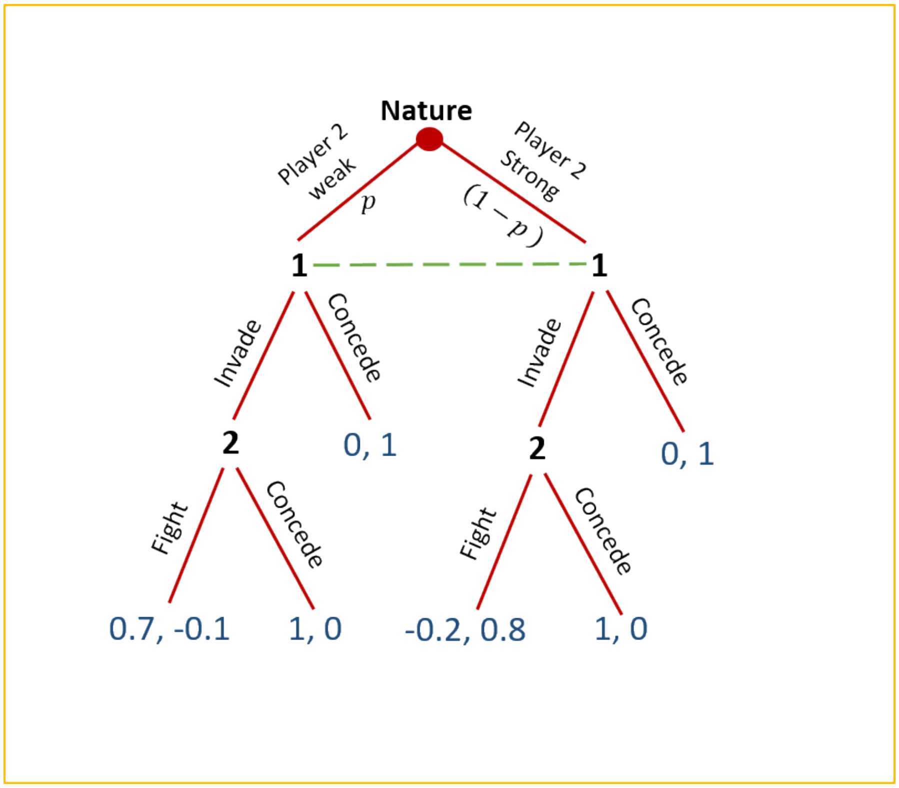
ESCALATION GAME WITH INCOMPLETE INFORMATION

Spaniel (2011) proposes a tweak to the Escalation Game by endowing Player 1 with less information than Player 2. We assume that Player 1 does not know for certain whether Player 2 is a "weak" or "strong" type. Player 1 therefore assigns probability p to Player 2 being weak and $(1-p)$ to Player 2 being strong. "Nature" has pre-assigned Player 2 his type, which Player 2 alone is aware of with certainty. Player 1 moves first. When he moves, Player 2 knows both his type and the move made by Player 1 in the first stage.[11] The decision tree for this game looks like this:

9. Note that depicting the game in extensive form makes it easy to identify the number of different subgames. In this case there are four subgames.

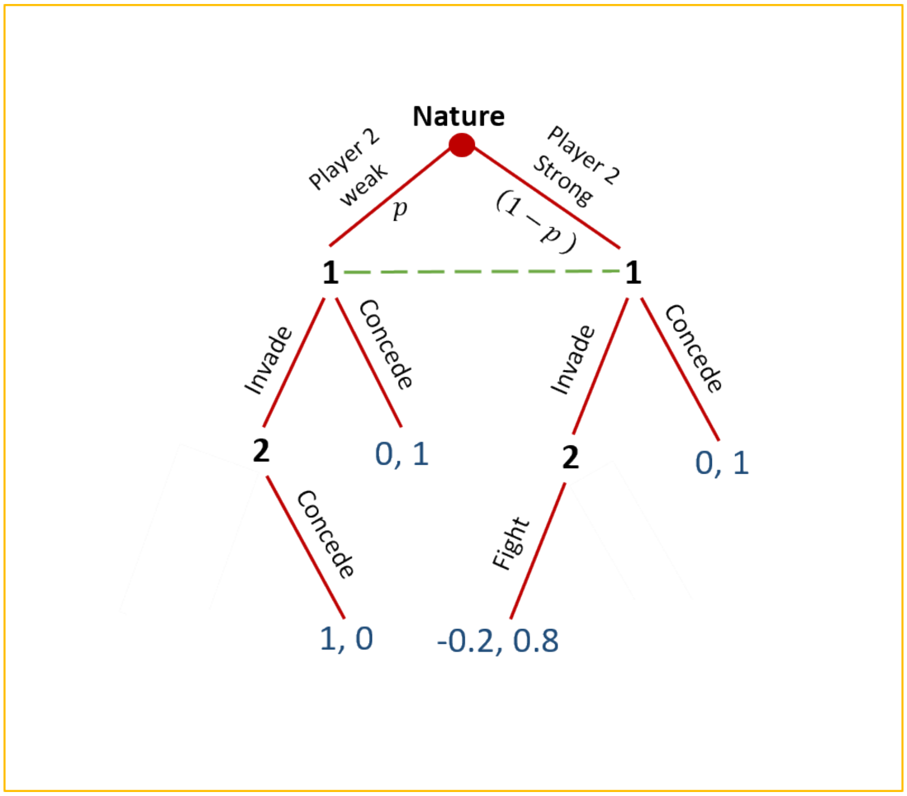
10. Here's a test to see how well you've grasped the logic behind the game's analytical equilibrium. What equilibrium would result if instead of having to pay \$2 by choosing to "concede" in the second stage, Player 2's required payment was instead something less than \$1?

11. This type of game is also known as a "screening game," where the lesser-informed player moves first.



In the initial stage (which we will call Stage 0), Nature determines whether Player 2 is weak or strong. Player 1 assumes Player 2 will be weak with probability p and strong with probability $(1-p)$, where $0 \leq p \leq 1$. As indicated by the green hashed line, when Player 1 moves in the first stage, she does not know for certain whether Player 2 has been determined as weak or strong. If it turns out that Player 2 was determined by Nature to be weak and Player 1 "concedes," the game ends with Player 1 receiving a payoff of \$0 and Player 2 receiving a payoff of \$1. If, on the other hand, Player 1 chooses to "invade," then Player 2 chooses between "fight" and "concede" in the second stage, resulting in payoffs of \$0.70 and -\$0.01 and \$1 and \$0, respectively for Players 1 and 2. If, instead, it turns out that Player 2 was determined by Nature to be strong and Player 1 chooses to concede in the first stage, the game ends with Player 1 again receiving a payoff of \$0 and Player 2 receiving a payoff of \$1. If Player 1 chooses to invade, then again, Player 2 chooses between "fight" and "concede" in the second stage resulting in payoffs of -\$0.20 and -\$0.08 and \$1 and \$0, respectively to Players 1 and 2. Whew!

Before working through the logic of the analytical equilibrium, notice that if and when the players reach the game's second stage, Player 2 will never choose to fight if he was determined by Nature to be weak (remember that Player 2 knows for certain whether he is weak or strong before play begins with Player 1). This is because the payoff from conceding at that stage is \$0, which is larger than the payoff of -\$0.1 associated with choosing to fight. Similarly, if Player 2 was determined by Nature to be strong, then he will never choose to concede if and when the players reach Stage 2 (\$0.8 > \$0). Thus, the decision tree for this game can now be depicted as the following:



Now, how do we solve for the game's analytical equilibrium?[12]

Here, Player 2 applies backward induction to find what's known as a Perfect Bayesian Equilibrium (PBE). As we already know, if Player 2 is the weak type and Player 1 has chosen to invade, then Player 2 should concede. If he is the strong type, then Player 2 should fight. We also know that Player 1 recognizes that she gets a payoff of $0 if she concedes in the first round, regardless of Player 2's type. If she instead chooses to invade in the first round, then Player 1's expected payoff from invading is $p - 0.2(1-p) = 1.2p - 0.2$. *This is merely the weighted average of Player 1's expected payoff when Player 2 is weak and her expected payoff when Player 2 is strong. Thus, invade is a better strategy than concede for Player 1 when* $1.2p - 0.2 > 0 \Longrightarrow p > 1/6$. *In other words, if the probability that Player 1 assigns to Player 2 being weak is greater than one-sixth, Player 1 should choose to invade in the first round. Otherwise, Player 1 should concede and be done with it.*

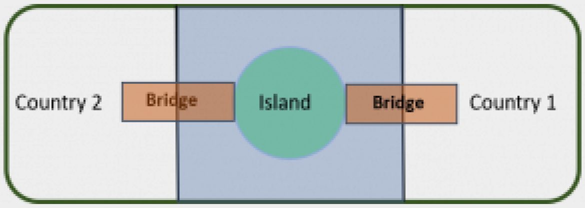
What's the outcome when you and your classmates play this more complicated version of the Escalation Game?

BURNING BRIDGES GAME

This game shares starkly similar features with the Escalation Game, but there is no uncertainty (thus, the analytical equilibrium is an SPE rather than a PBE). The SPE has much to say about the relationship between two tenacious competitors. Spaniel (2011) portrays the game as follows:

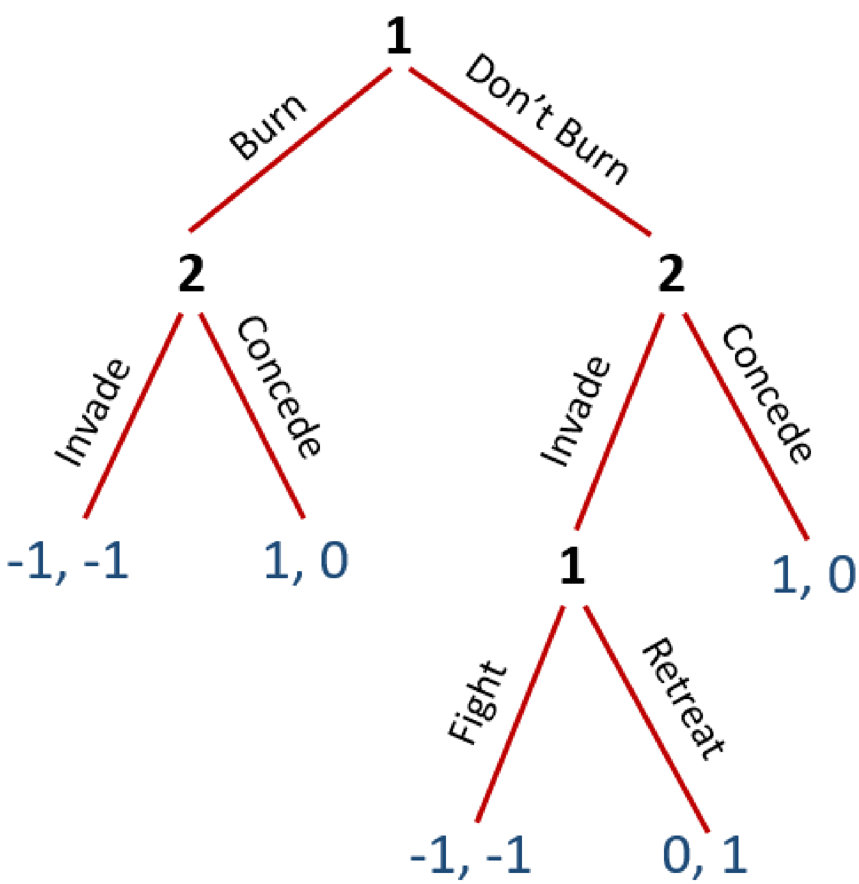
12. This equilibrium is known as a Perfect Bayesian Equilibrium (PBE) rather than an SPE because of the uncertainty that at least one of the players is forced to contend with. Similar to Nash, Thomas Bayes is considered a towering figure. He was an 18th-century English statistician, philosopher, and Presbyterian minister who is known for formulating a specific case of the theorem that bears his name: Bayes Theorem. Bayes never published his theory himself—his notes were edited and published posthumously.

Suppose an island is located between two countries. Each country has a bridge to the island.



Country 1 decides to cross over its bridge to the island in an act of war. Country 1 must then choose whether to burn the bridge behind it or not.

The game's structure is depicted by the following decision tree:



Recall that this game starts with Country 1 already having crossed the bridge onto the island. Country 1's choice in the first stage of the game is, therefore, whether or not to burn the bridge behind it. If Country 1 burns the bridge, then Country 2 must decide whether to cross its bridge and invade the island as well or to concede the island to Country 1. The resulting payoffs for the two countries are as shown. If, instead, Country 1 chooses not to burn its bridge, then if Country 2 also decides to invade the island, Country 1 must then choose whether to stand and fight or retreat back over its bridge to safety. Otherwise, if Country 2 decides to concede, then the result is the same as when Country 1 decides to burn its bridge.

The logic for the analytical equilibrium goes like this:

> *If Country 1 chooses not to burn its bridge, Country 2 will choose to invade the island knowing that Country 1 will then choose to retreat (since \$0 > -\$1), thus giving Country 2 a payoff of \$1, which is larger than the payoff it would have gotten had it instead chosen to concede. Thus, Country 1's payoff from choosing not to burn its bridge is ultimately \$0. If Country 1 instead chooses to burn its bridge, Country 2 will choose to concede (since \$0 > -\$1), giving Country 1 a payoff of \$1. Thus, Country 1 choosing to burn its bridge and Country 2 responding by conceding the island to Country 1 is this game's SPE.*

There are historical examples of this game having been played between civilizations and countries and even individuals. For example, Collins (1989) recounts an incident in 711 AD when Muslim forces invaded the Iberian Peninsula, and commander Tariq bin Ziyad ordered his ships to be burned, thus signaling to his troops that they had passed the point of no return. Harvey (1925) recounts a similar incident in Myanmar (formerly Burma). In the Battle of Naungyo, during the Toungoo-Hanthawaddy War in 1538, the Toungoo armies led by commander Kyawhtin Nawrahta (later known as Bayinnaung) faced the superior force of Hanthawaddy on the other side of a river. After crossing the river on a makeshift bridge, Bayinnaung ordered the bridge to be destroyed. Similar to Muslim commander bin Ziyad, Bayinnaung took this action to spur his troops forward in battle and provide a clear signal that there would be no retreat. In both cases, the commanders were victorious.

Have you ever burned your proverbial bridge in negotiations with an employer, a friend, or maybe even a family member? If the answer is "yes," chances are you are not alone. Most *Homo sapiens*, if they live long enough, are eventually confronted with having to play a game like this. Typically, it requires a curious mixture of courage and desperation for a player (in our case a Country 1) to summon the will necessary to achieve the game's analytical equilibrium.

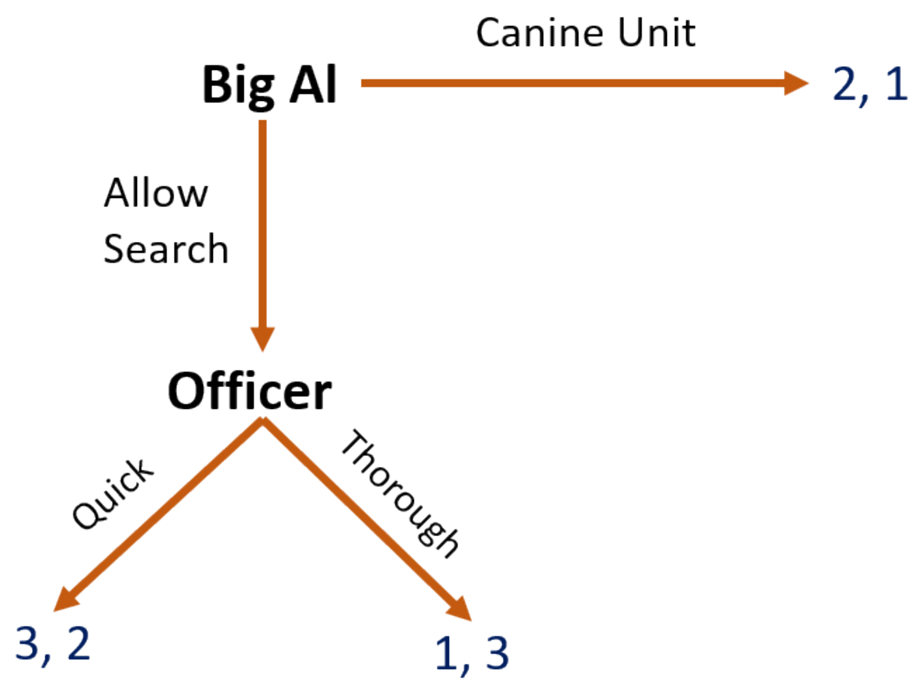
POLICE SEARCH

Spaniel (2011) describes a game he once remembers having played himself. The title of the game says it all:

> Suppose a police officer pulls Big Al over and asks to search his vehicle. Big Al can let the police officer search the vehicle (which could be a quick or a thorough search, depending upon the police officer's preferences) or refuse and force the officer to call in the Canine Unit. Big Al's preferences are **Quick Search** $\succ$ **Canine Unit** $\succ$ **Thorough Search**, while the police officer's preferences are **Quick Search** $\succ$ **Thorough Search** $\succ$ **Canine Unit**.
>
> Without actually knowing Big Al's preferences, the officer nevertheless claims that "a Quick Search is more preferred for both of us than calling in the Canine Unit."

Recall from Chapter 3 that the symbol $\succ$ stands for "strictly preferred to." Thus, we can say that based upon the information given above, Big Al strictly prefers a quick search as opposed to summoning the canine unit, and strictly prefers the canine unit as opposed to a thorough search. In contrast, the police officer strictly prefers the thorough search over a quick search, and strictly prefers the quick search over bringing in the canine unit.

It helps if this game is depicted as a decision tree.



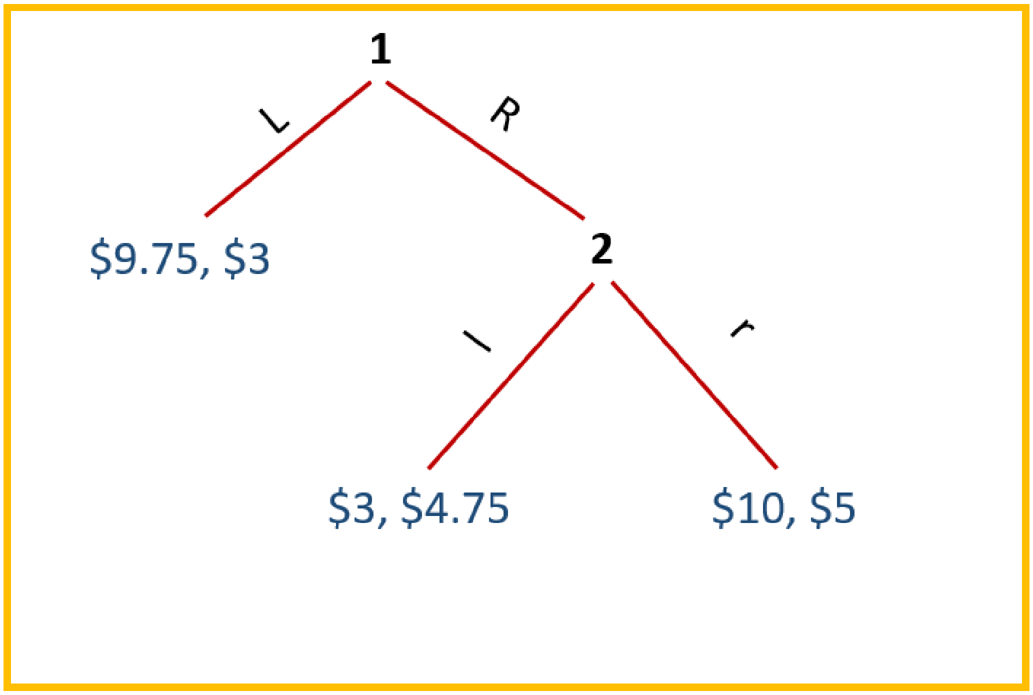
Note that the hypothetical payoffs associated with each choice in the decision tree correspond with Big Al's and the police officer's respective preference rankings. To find this game's analytical equilibrium, we need to assume common knowledge among Big Al and the officer (i.e., each player knows both his own payoffs and those of the other player). Common knowledge has been implicitly assumed for each of the games examined thus far except, of course, in the case of the Escalation Game with Incomplete Information. The logic for the analytical equilibrium is as follows:

> *If Big Al allows a search, the officer will choose to do a thorough search, implying Big Al's payoff is 1 and the officer's is 3. If, instead, Big Al does not allow the search, the Canine Unit is called in, resulting in payoffs to Big Al and the officer of 2 and 1, respectively. Clearly, the game's SPE is Big Al not allowing a search, and the officer calling in the canine unit.*

In some sense, by holding his ground on not allowing the officer to search his car, Big Al is burning his bridge with the officer. The equilibrium outcome is driven by the fact that the officer cannot credibly commit to conducting a quick search if Big Al were to allow the officer to conduct a search. Sadly, the resulting SPE for this game is inefficient since, as the officer originally pointed out, both Big Al and the officer prefer the quick search. If Big Al and the officer were *Homo economicus* rather than *Homo sapiens*, they would mutually trust each other in this particular context, and the quick search would be conducted. Both the officer and Big Al would save valuable time, and the canine unit would get more rest.

TWO-STAGE ITERATED DOMINANCE GAME

Beard and Beil (1994) propose the following game:



Player 1 chooses first, moving either left (L) or right (R). If she moves L, the game ends with Players 1 and 2 receiving payouts of $9.75 and $3, respectively. If, instead, Player 1 moves R, then Player 2 chooses in the second stage whether to move left (l) or right (r). The payoffs to both players are then as given. The game's SPE is determined via the following logic:

> *By backward induction, Player 1 considers what Player 2 will choose if the game proceeds to the second stage. Player 2 will choose r since $5 > $4.75. This results in $10 for Player 1, which is larger than the $9.75 payout she would obtain if she decides to move L in the first stage. Thus, Player 1 choosing R and Player 2 choosing r is the game's SPE (denoted (R,r)).*

After playing this game with various groups of *Homo sapiens*, Beard and Beil (1994) report that 66% of Player 1s chose to move L.[13] In the 34% of instances where Player 1s moved R, their choices were met with Player 2's self-interested response of r 83% of the time. Beard and Beil calculated Player 1's faith in Player 2's rationality required to justify choosing R in the first stage (which they label a threshold probability $p(r \mid R)$ as equaling 0.97. In other words, Player 1s reported needing to believe that Player 2 would choose r in the second stage 97% of the time before they could justify choosing R in the first stage. Since Player 2s chose r only 83% of the time, the threshold was not quite met on average.

DIRTY FACES

Littlewood (1953) invented this iterated-knowledge game whereby three ladies, A, B, and C, in a railway carriage all have dirty faces and are all laughing. Because none of the ladies can see their own face to know for certain whether their face is dirty, they must infer from the laughter of the other two ladies whether their own face is dirty. The version of this game presented in Camerer (2003) involves only two players, but the notion of iterated knowledge is nonetheless retained.

13. These results are for Beard and Beil's (1994) baseline treatment. The authors considered several other treatments where the payoffs for the two players were modified. The results for most of these alternative treatments were qualitatively similar to those obtained in the baseline treatment.

Two players have independently and randomly drawn their "types", either X or O, with probabilities of 80% and 20%, respectively. After observing the other player's type—but not their own type—the two players choose either "Up" or "Down." Payoffs for each player are given in the matrix below.

	Type: X (0.8)	Type: O (0.2)
Up	0	0
Down	5	-10

Thus, if a player chooses Up, he earns nothing. If a player chooses Down, he earns $5 if he is type X and loses $10 if he is type O. When at least one player is type X, both players are told, "At least one player is type X." Successive rounds of the game are played (with each player retaining their original type) until one of the players chooses Down. After each round, the players are told of the other player's choice.

The logic for the analytical equilibrium goes like this:

> *There are two cases to consider—the XO case (one player is X and the other is O) and XX (both are X). We do not need to consider the OO case since, when this happens, each player will know immediately that he is an O type (how?) and thus, neither of the players will ever choose Down.*[14] *In the XO case, the player who is X can infer this fact (how?). He then moves Down. In the XX case, both players know there is at least one type X (after the announcement is made that at least one of the players is type X), and they know the other player is X, but they still know nothing for certain about their own type. Each player, therefore, chooses Up in the first round and is then told of the other player's choice. Player 1, for example, is told that Player 2 chose Up. Player 1, therefore, infers that Player 2 must have known Player 1 was a type X. Otherwise Player 2 would have chosen Down. Thus, Player 1 infers his own type from Player 2's behavior—he must be an X. Hence, Player 1 chooses Down in the second round. And therefore, for the case of XO, we expect the type X player to choose Down in the first round of the game, while in the case of XX, we expect both players to choose Down in the second round of play (after the first announcement of player choices has been made by the experimenter).*

Weber (2001) enlisted a small group of participants to play the Dirty Faces game. Recall that in Round 1, we expect the equilibrium to be (Down, Up) when the players are types X and O, respectively, and (Up, Up) when both players are type X. In Round 2, played only by players who are both X types, we expect the (Down, Down) equilibrium to result. The author found that in the XO case, player pairs behaved like *Homo economicus* seven out of eight times across two different trials by choosing Down when they were type X. In the trickier XX case, players are predicted to choose (Up, Up) in the first

14. In actual experiments conducted by Weber (2001) with Homo sapiens, results for cases where both players drew the O type are left unreported.

round, followed by (Down, Down) in the second round (after each player is informed of the other player's choice). In Weber's experiment, player pairs chose (Up, Up) in the first round 14 out of 18 times. But, only four of the 14 player pairs who chose (Up, Up) in the first round chose (Down, Down) in the second round.

The evidence for *Homo sapiens* is, therefore, mixed. They seem to mimic *Homo economicus* in the XO case rather well, but not nearly as well in the XX case.

TRUST GAME

Consider the following game proposed by Berg et al. (1995):

> An Investor has \$$x$ which she can keep or invest. Suppose she decides to invest \$$T$ and keep \$($x - T$). The investment of \$$T$ earns a return, at a rate of (1 + r), and becomes \$(1 + r)T. Another player, the Trustee, must now decide how to share the new amount \$(1 + r)T with the Investor. Suppose the Trustee decides to keep \$$y$ and thus returns \$[(1 + r)T – y] to the Investor, resulting in a payoff of \$$y$ for the Trustee and \$[x – T + (1 + r)T – y] = \$[x – y + rT] for the Investor.
>
> Thus, \$$T$ is a measure of **trust** and \$[(1 + r)T – y] is a measure of **trustworthiness**.
>
> For our game, let x = \$200 and r = 1.

Despite the relatively complicated calculations involved in determining the returns to the Investor and Trustee for different possible investment and share values, the logic behind the game's analytical equilibrium is a straightforward application of iterated dominance practiced by the investor, in particular backward induction.

> *Because the Investor anticipates that the Trustee will keep whatever investment is made, the Investor chooses to keep the entire \$200 and thus invests nothing! Consequently, in this game's SPE, there is no trust displayed by the Investor and no opportunity for trustworthiness, or "direct reciprocity," to be displayed by the Trustee.*

Not so with *Homo sapiens*. Berg et al. find that Investors invest roughly 50% on average (i.e., T/x = 50%), and Trustees repay roughly 95% of what was invested (i.e., $(((1+r)T-y))/T$ = 95%), which equals a negative return to trust and a correspondingly slight lack of trustworthiness![15] In a modest tweak, Buchan et al. (2000) engage Asian and American subjects in a trust game where the Investor knows she will receive the return from a different (i.e., third-party) Trustee rather than the original Trustee. The authors find that both trust and trustworthiness decrease relative to Berg et al.'s findings (i.e., a sense of karma—that one would hope exists among Trustees—does not restore trust and trustworthiness). Buchan et al.'s (2000) results are contained in the table below.

15. The threshold for displaying trustworthiness in this game is 100% of the Investor's investment returned by the Trustee.

Player Groups	Pair	Foursome	Society	Overall
Percent Invested (i.e., Trust Level)				
American – Chinese	76	49	49	54
Japanese – Korean	51	48	28	41
Mean	64	48	39	47
Percent of Tripled Investment Returned (i.e., Trustworthiness Level)				
American – Japanese	28	13	11	15
Chinese – Korean	41	25	18	25
Mean	35	19	15	20

(Camerer 2003)

The different groups are similar (e.g., when it comes to Investors, American and Chinese subjects behaved similarly to each other, and Japanese and Korean subjects behaved similarly as well). As Trustees, American and Japanese subjects behaved similarly, and Chinese and Korean subjects behaved similarly. "Pair" represents a control treatment where the Investor receives a return from the same Trustee she invests with. The Foursome treatment refers to a version of the game where there are two Investors (A and B) and two Trustees (C and D). Investor A originally invested with Trustee C but is repaid by Trustee D, and Investor B originally invested with Trustee D but is repaid by Trustee C. Investors A and B know this "cross repayment" is occurring. The Society treatment refers to the case where Investors and Trustees are in separate rooms, and which Trustee repays which Investor is determined randomly. Investors A and B, therefore, do not know which Trustee has been assigned to them ahead of time. The "Overall" column provides the average across these different treatments.

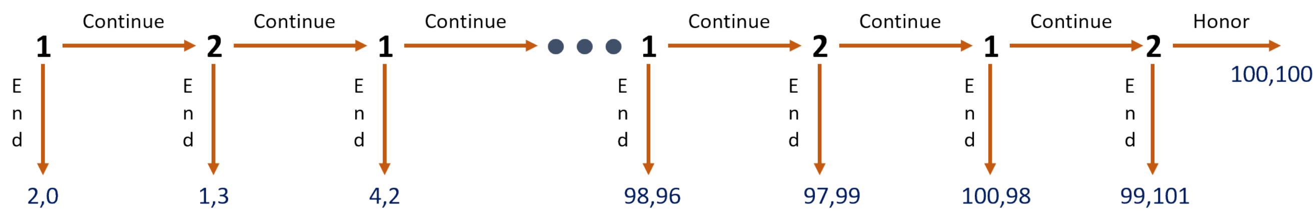
We see that, on average, Investors chose to invest 64% of their initial amount with the Trustee in the control treatment (which exceeds Berg et al.'s finding of 50%). However, alternative pairings of Investors with Trustees lead to a reduction in the investment made by investors (i.e., a decrease in trust). Overall, trust is effectively displayed at a 47% level. Similarly, although in the control treatment 105% of the Investor's initial investment is returned (35 x 3 = 105%), the overall return on investment is only 60% (20 x 3 = 60%), which represents a markedly lower level of trustworthiness than found by Berg et al (1995).

Carter and Castillo (2011) conducted similar trust experiments with relatively poor and lower-educated residents in rural and urban communities in South Africa. On average, Investors trusted their anonymous partners with 53% of their investable income, remarkably close to the percentages observed in the U.S. experiments performed by Berg et al. (1995). However, the amounts invested varied substantially depending upon which village the Investor was from. Similarly, on average, Trustees reciprocated in a trustworthy manner by returning 100% of the Investor's investment. Carter and Castillo also found that trust and trustworthiness went hand-in-hand—residents located in villages with higher levels of trust also tended to exhibit higher levels of trustworthiness. Further, in urban communities, higher levels of trust and trustworthiness are correlated with higher levels of household expenditures (a proxy for household well-being), while in rural communities this relationship is reversed—higher levels of trust and trustworthiness are associated with lower levels of household expenditure.

One potential explanation for these latter results is that trust in a rural village is prone to moral hazard (Just, 2013). Moral hazard occurs when a person's actions are not fully observed by others, yet they can affect the welfare of both that person and others. A trusting rural village might mean that residents generally assume that everyone else will perform their civic and economic duties and therefore do not need to be closely monitored. The marginal gain in household well-being from trusting others a bit more is, therefore, relatively small. In a less-trusting village, residents will monitor each other more closely to learn whether others are actually performing their work, potentially leading to a larger marginal gain in household well-being as a result of being able to trust others more. Alternatively, very few residents in an urban area are related to each other. This means that trusting others more might allow you to build a wider social network, potentially creating some substantial gains in household well-being from being able to trust others more.[16], [17]

MULTI-STATE ("CENTIPEDE") TRUST GAME

The multi-stage trust game is best represented in the form of a decision tree:



Player 1 acts as the investor in the first stage by choosing whether to end the game immediately by not investing (in which case she obtains a payoff of \$2 and Player 2—acting as the Trustee in this stage—gets nothing), or by investing and continuing the game to the second stage. In the second stage, Player 2 now acts as the Investor and decides whether to end the game by not investing (in which case she obtains \$3 and Player 1—acting as the Trustee in this round—gets \$1), or by investing and continuing the game to the third stage where, once again, Player 1 acts as the Investor and Player 2 the Trustee. As depicted, the game can be played up to 100 stages. And now we see how this game got its moniker—its decision tree bears a striking resemblance to a centipede.

The logic for the analytical equilibrium goes like this:

16. Stanley et al. (2011) conducted trust games in the US with participants of different races and found that differences in trust and trustworthiness can be partially explained by differences in implicit attitudes toward race. Similar differences in trust and trustworthiness between races were discovered in earlier experiments conducted by Glaeser et al. (2000). In their laboratory experiments, Johansson-Stenman et al. (2009) find weak differences in trust and trustworthiness among Bangladeshi individuals with different religious identities (Hindu and Muslim). Croson and Buchan (1999) find no significant effect of gender on Investors' level of trust. However, the authors find that women Trustees reciprocate significantly more of their wealth in trust games than men, both in the US and internationally. In a novel laboratory experiment, DeBruine (2002) finds that trust game participants were more likely to trust others who look more like they do, which suggests that people over generations evolve toward promoting the well-being and survival of others with whom they share a resemblance.

17. Carter and Castillo (2004) conducted a trust experiment with survivors of Hurricane Mitch, which devastated rural Honduran communities in 1998, with the goal of measuring the extent to which trust among community members helped spur recovery efforts. As the authors point out, while many communities received some inflow of external aid, the absence of insurance contracts and thinness of capital markets meant that most households had to rely either on their own resources to muster an economic recovery, or on resources that they could broker through social relationships. Econometric analysis of the experimental data provided evidence of durable community norms, such as trust that is reinforced by social interactions. The analysis shows that trust played a strong, but uneven role in facilitating recovery from Hurricane Mitch, assisting most strongly a favored subset of households. While establishing the importance of norms such as trust, Carter and Castillo's analysis warns against the presumption that all community members are equally well-served by the social mechanism of trust in the face of recovery from a natural disaster.

Via backward induction, Player 2 (as Investor) ends the game in the final stage with no investment (thus earning a payoff of $101). Knowing this will happen, Player 1 (as Investor) ends the game in the penultimate stage with no investment (thus earning a payoff of $100, which is larger than the $99 she would have earned as Trustee had the game instead ended in the final stage). As the game unravels in mistrust all the way back to the first stage, Player 1 (as Investor) chooses to end the game in the first stage with no investment (thus earning a payoff of only $2). This game's SPE is woefully inefficient. Too bad Homo economicus are so self-interested. Had they been able to cooperate with each other they could have played to the final round, with Player 1 earning $99 and Player 2 earning $101.

Camerer's (2003) survey of the literature suggests that *Homo sapiens* tend to reciprocate trust and trustworthiness for a few stages before one of the players ends the game. This may be a case of Player 1 believing Player 2 lacks common sense. Player 1 thus plays Continue in the first stage, sending a signal that she trusts Player 2, who then also chooses Continue in the second stage. In cases where altruistic players are involved, the Honor payoffs in the final period may be interpreted as (101, 102), and the players are self- and dual-motivated to advance all the way to the final stage.

In a novel laboratory experiment, Scharlemann et al. (2001) led participants to believe that they were playing the Centipede game with a randomly paired partner. Before choosing a strategy, Player 1 was given a photograph of the player (Player 2) to whom he was purportedly paired, and likewise for Player 2, who was given a photograph of purported Player 1. In reality, both players were playing against predetermined strategies programmed into a computer. Nevertheless, each player believed he was playing against the player in the picture.

There were actually two pictures of each purported player. One of the photos depicted the player smiling, and the other depicted the player with a straight face. Participants were randomly assigned to see either a smiling or a straight-faced partner. Overall, the authors found that participants were roughly 13% more likely to choose Continue at the first stage when their supposed partner was smiling in the photograph. Thus, there is some evidence to suggest that participants interpreted a smile on their partner's face as signaling trustworthiness.

Further, this "smile effect" was noticeably larger for male participants than for female participants. Male participants trusted a smiling partner roughly 20% more than a non-smiling partner, while female participants trusted smiling partners only 6% more. Additional experiments by the authors using other facial expressions also had an impact on the willingness of participants to trust each other.

As the saying goes, what's in a smile? Perhaps the trust it inspires in those who are graced by one.

To wrap up our exploration of the centipede game, consider McKelvey and Palfrey's (1992) version where, rather than the payoffs associated with successive stages of the game alternating between increases and decreases (as depicted in the game above), the payoffs increase at a constant rate for both players. Note that investigating possible effects associated with changes in the way payoffs evolve in the centipede game is reminiscent of investigations into the effects of changing the payoffs (or stakes) in the Ultimatum Bargaining and Beauty Contest games discussed earlier (recall that changing the stakes in these two games generally had no impact on the games' respective outcomes in experiments with *Homo sapiens*).

In their laboratory games, McKelvey and Palfrey start with a total pot of $0.50 divided into two smaller pots of $0.40 and $0.10. Each time a player chooses to "pass" (i.e., continue), both pots of money are multiplied by two. The authors construct both a two-round (four-move) and a three-round (six-move) version of the game. McKelvey and Palfrey also consider a version of the four-move game in which all payoffs are quadrupled. The authors found that, as with the traditional centipede game described above, the SPEs for these two versions of the game are for Player 1 to "take" (i.e. end) the games in the first round.

In each experimental session, McKelvey and Palfrey use a total of twenty subjects, none of whom had previously played a centipede game. The subjects (students from Pasadena Community College and the California Institute of Technology), were divided into two groups at the beginning of the session, called the Red and Blue groups. In each game, the Red player was the first mover, and the Blue player was the second mover. Each subject then participated in ten games, one with each of the subjects in the other group. Subjects did not communicate with each other except through the choices they made during the game. Before each game, each subject was matched with another subject of the opposite color with whom they had not yet been previously matched. The paired subjects then played either the four-move or six-move game.

McKelvey and Palfrey found that in only 7% of the four-move games, 1% of the six-move games, and 15% of the high-payoff four-move games did Player 1 choose "take" in the first round. The subjects do not iteratively eliminate dominated strategies as they would if, like *Homo economicus*, they played SPE strategies.[18] In each of the sessions, the probability of a player choosing "take" increases as they get closer to the game's last move. Thus, as subjects gain more experience with the game, their behavior mimics that of *Homo economicus*.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013), those marked with a "‡" are adopted from Cartwright (2014), and those marked with a "♉" are adopted from Dixit and Nalebuff (1991).

1. Recall the Ultimatum Bargaining game studied in this chapter. A Proposer makes an initial offer to a Responder of how to split $100. If the Responder accepts the Proposer's offer, the $100 is split accordingly. If the Responder rejects the Proposer's offer then both receive $0. As we showed, the analytical equilibrium for this game is the Responder offering $0.01 and the Responder accepting. How would the analytical equilibrium for this game change if the game instead adhered to the following rules: In the **first stage**, the Proposer makes an offer to the Responder of how to split the $100. In the **second stage**, the Responder can choose to either accept the offer as is, or agree to a flip of a fair coin. If the coin comes up "Heads," then the game moves to the third stage. If the coin comes up "Tails," the game ends with Proposer and Responder each receiving $0. In the **third stage**, the Proposer can decide to either split the $100 50%-50% (i.e., give $50 to the Responder and keep the remaining $50), or agree to a flip of a fair coin. If the coin comes up "Heads," the Proposer keeps $75 and gives the Responder $25. If the coin comes up tails, the Proposer instead gives the Responder $75 and keeps $25. What is the analytical equilibrium for this version of the Ultimatum Bargaining game? Explain.

18. Instead, the players exhibit what on the surface appears to be altruistic behavior. However, as McKelvey and Palfrey point out, if a selfish player believes that there is some likelihood that each of the other players may be altruistic, then it can pay the selfish player to mimic the behavior of an altruist in an attempt to develop a reputation for choosing to "pass." The authors surmise that the incentives to mimic are very powerful, in the sense that a very small belief that altruists are in the subject pool can generate a lot of mimicking, even when the players face a very short time horizon. McKelvey and Palfrey ultimately estimate that their sample consists of only 5% altruists.

2. Knowing what you know about basins of attraction, or path dependence, explain your strategy for playing repeated rounds of the Continental Divide game.

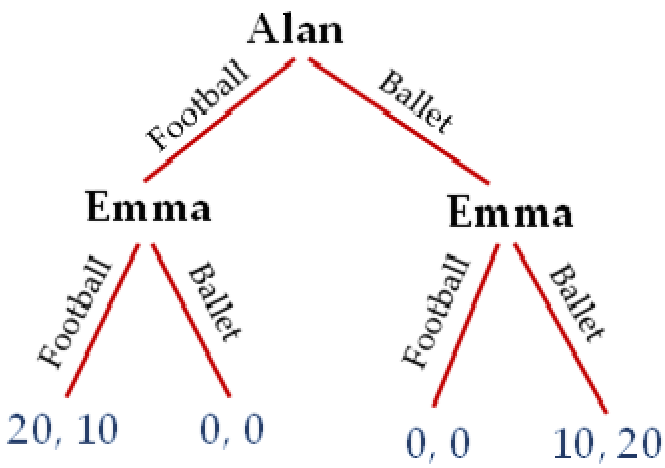
3. What would the analytical equilibrium of the Beauty Contest game be if factor p were instead set equal to a number greater than 1, say, 1.4? Show how you arrive at your answer.

4. Recall the Traveler's Dilemma game where two travelers simultaneously submit claims to the airline for their lost luggage ranging between $300 and $750. The airline pays both travelers the minimum claim, and then subtracts $50 from that amount for the player who submitted the higher of the two bids and adds $50 to that amount for the player who submitted the minimum of the two bids. In comparison with the analytical equilibrium for this game, explain why the airline should expect to pay out more in claims to two *Homo economicus* travelers and less in claims to two *Homo sapiens* travelers if it changed the game accordingly: The travelers get to choose one of two options. Option 1 is the same as the original game, except now the lower-bound on the range of claims is $250 instead of $300. For Option 2, the airline flips a fair coin. If the coin comes up "Heads," the traveler receives $750; if it comes up "Tails," the traveler receives $0.

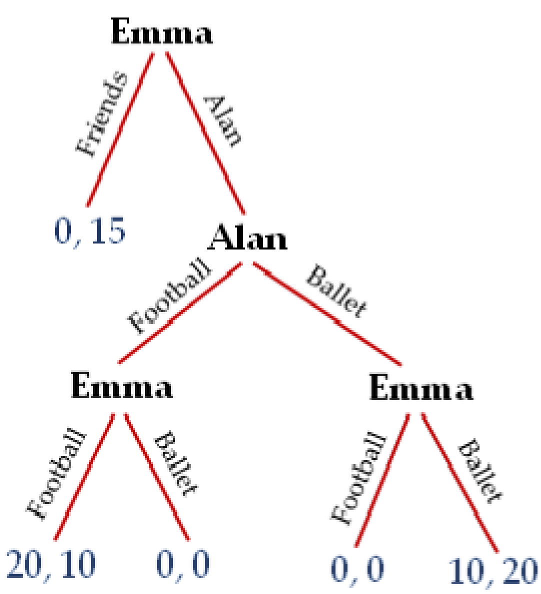
5. Calculate the Perfect Bayesian Equilibrium (PBE) for the Burning Bridges game if Country 2 is uncertain as to whether Country 1 has burned its bridge after it (Country 1) has occupied the island.

6. † Suppose you are a bank manager. You know that if depositors trust your bank, they will be willing to accept a lower interest rate on deposits. (a) Given what we know of how people develop trust, what might you do to enhance your depositors' trust? (b) What steps might you take to ensure that your loan officers can avoid potential pitfalls when it comes to originating loans to business owners and other customers?

7. ‡ Suppose Alan and Emma are locked in a sequential-move version of the Battle of The Sexes game depicted below (you will be introduced to the classic simultaneous move version of this game in Chapter 8). Alan chooses first, choosing to attend either the football game or ballet. Next, Emma chooses the football game or the ballet. The first value(s) at each terminal node of the decision tree represents the payoff(s) accruing to Alan, and the second value(s) represents the payoff(s) accruing to Emma (all in dollars). What is the most likely outcome of this game? Discuss how you have arrived at your answer.



8. ‡ Now suppose that in her sequential Battle-of-The-Sexes game with Alan, Emma has an outside option that comes into play at the outset. She chooses to either go out with her friends or "throw her lot" in with Alan for either the football game or ballet. What is the most likely outcome of the game now? Again, discuss how you have arrived at your answer.



9. Choose a sequential game you learned about in this chapter. In what way might permitting communication between the players before the game begins affect the game's outcome?

10. Consumers who choose to purchase used vehicles from used-vehicle salespeople complain that the bargaining process resembles an ultimatum bargaining game. Why might this be the case?

11. ◊ In the game of roulette, betting is based on where a ball will land when a spinning wheel

stops. In the game's simplest form, there are numbers zero through 36 on the wheel. When the ball lands on zero, the players win nothing (or, alternatively stated, the "house" wins). The safest bet in roulette is simply to bet that the wheel will stop on an even or odd number, with numbers zero through 36 on the wheel the chances of winning are 18/37, or a little less than 49%. This bet pays "even money" (i.e., a $1 bet returns $1, leaving the player with a total of $2). A second possible bet would be that the wheel will stop spinning on a multiple of three (e.g., the numbers 3, 6, 9, etc.) for a 12/37, or slightly larger than a 32% chance of winning. This bet pays "two-to-one" (i.e., a $1 bet returns $2 for a total of $3). When players place their bets ahead of the spin of the wheel, they inevitably do so sequentially – no rule says they must place their bets simultaneously with the spin of the wheel or with each other. Whenever a player bets wrong, e.g., places an even-money bet on an odd number and the wheel stops on an even number, or vice-versa, the player loses whatever amount of money was placed on the bet. Suppose that in this game the only possible bets are (1) the even-money bet on an even or odd number, or (2) the two-to-one bet on a multiple of three. Bonnie and Clyde are down to the last spin of the wheel. Whoever has amassed the most money (in terms of the value of their chips) at end of the final spin buys dinner. Bonnie has $700 worth of chips and Clyde has only $300. All else equal, what is Clyde's best bet? What is Bonnie's best bet? Does either player have a first-mover advantage?

12. Suppose a new store called *Newbies* is considering entering a market that is currently dominated by a store called *Oldies* (i.e., *Oldies* is currently a monopoly in this market, earning $200,000). It is common knowledge that if *Newbies* enters the market and *Oldies* accommodates (i.e., does not wage a price war), *Newbies* will earn $100,000 and *Oldies* will also earn $100,000. If *Oldies* instead chooses to launch a price war, then *Newbies* will ultimately lose $200,000 and *Oldies* will lose $100,000. Draw the decision tree for this game and determine its subgame equilibrium.

13. Suppose the state of Utah institutes a new statewide program called the *Utah Brigades*, which requires every high school senior to register for a year of public service to the state upon graduation. Worried that this new requirement may lead to mass civil unrest among the students, and unwilling to punish each student who refuses to register, the state announces it will go after evaders in alphabetical order by last name. (a) Explain why this approach could lead to full compliance with the registration. (b) Would this approach still work if the state announced it will go after evaders by Social Security Number, in either ascending or descending order?

14. Suppose two parents would like their adult children to communicate with them on a more regular basis. They announce a new quota that each respective child must meet in order to receive their portion of the parents' inheritance: one visit and two phone calls per week. Any child who does not meet the quota on any given week is disinherited, and the remaining children split the inheritance among themselves. Recognizing that their parents are very unlikely to disinherit all of them, the children get together and agree to cut back on their visits and phone calls, potentially down to zero. What change could the parents make to their will to

ensure that the children meet their quota?

Media Attributions

- Continental Divide Matrix © Van Huyck et al.; 1997 Elsevier Science B.V. is licensed under a All Rights Reserved license
- Figure 7 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 8 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 9 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Burning Bridges Figure © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 11 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 13 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 14 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Dirty Faces Figure © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 1a (Chapter 7) © Colin F. Camerer; Princeton University Press is licensed under a All Rights Reserved license
- Figure 17 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 8 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 9 (Chapter 7) © Arthur Caplan is licensed under a CC BY (Attribution) license

CHAPTER 8.

SOME CLASSIC SIMULTANEOUS-MOVE GAMES

PRISONER'S DILEMMA

This is undoubtedly one of the most well-known simultaneous games *Homo sapiens* play (for the most part unwittingly) in their day-to-day lives. Prisoner's Dilemmas abound in our social interactions, in particular, governing how we manage natural resources collectively. You may have heard of the "tragedy of the commons" when it comes to managing fisheries, rangelands, local watersheds and airsheds, or global climate change. It turns out that a Prisoner's Dilemma is at the root of these types of resource management challenges. It is a dilemma that confronts us daily and drives individual decision-making in a social setting.

The game is presented below in its "strategic form"—a matrix containing the payoffs each of the two players will obtain from their respective choices when they move simultaneously as opposed to sequentially.[1] There is common knowledge in this game in the sense that Player 1 knows not only her payoffs listed in the matrix, but also Player 2's. Player 2 likewise knows not only his payoffs, but also Player 1's. The payoff matrix for a Prisoner's Dilemma game is depicted below:

		Player 2	
		Cooperate	Deviate
Player 1	Cooperate	4 , 4	1 , 6
	Deviate	6 , 1	2 , 2

In this game, both players simultaneously choose whether to Cooperate or Deviate. If both players choose Cooperate, then they both receive payoffs of $4 each. If both instead choose Deviate, then they receive payoffs of only $2 each. If Player 1 chooses Cooperate but Player 2 chooses Deviate, then Player 1 receives a payoff of only $1 while Player 2's payoff jumps to $6. Likewise, if Player 1 chooses Deviate when Player 2 chooses Cooperate, then Player 1's payoff jumps to $6 and Player 2's falls to $1.

Because moves in this game are made simultaneously, the solution concept is not SPE or BPE.

1. As we have learned thus far, games where players move sequentially are generally depicted as decision trees (or, in "extensive form"). In contrast, games where players move simultaneously are generally depicted as payoff matrices (or, in "strategic form").

Rather, it is either a "pure strategy equilibrium" or a "mixed strategy equilibrium." It turns out that the Prisoner's Dilemma solves for a unique pure strategy equilibrium (PSE). We will encounter simultaneous games that solve for mixed strategy equilibria (MSE) a bit later in this chapter.

The logic for this game's analytical PSE goes like this:

> *Both players consider their payoffs associated with Cooperate or Deviate given what the other player could decide to do and, in this way, devise their respective strategies. For instance, Player 1 first considers her payoffs when Player 2 chooses to Cooperate. Because \$6 > \$4, Player 1's best strategy is to Deviate when Player 2 chooses to Cooperate. Next, Player 1 considers her payoffs when Player 2 chooses to Deviate. Because \$2 > \$1, Player 1's best strategy is again to Deviate when Player 2 chooses to Deviate. Because Player 1's best strategy is to Deviate regardless of whether Player 2 chooses to Cooperate or Deviate, we say that Player 1 has a "dominant strategy" to choose Deviate no matter what Player 2 decides to do! Applying the same logic to Player 2's decision process, we see that Player 2 also has a dominant strategy to choose Deviate no matter what Player 1 decides to do. Thus, the PSE for this game is both players choosing to Deviate (i.e., (Deviate, Deviate)).*[2]

What a shame! Both players choose to Deviate, and, as a result, they attain payoffs of only \$2 each. This equilibrium is woefully inefficient. Had the two *Homo economicus* not been so self-interested and oh so rational, perhaps they could have agreed to Cooperate and earned \$4 each instead of just \$2. Such is the essence (and bane) of the Prisoner's Dilemma.[2]

Ironically (or should I say, sadly), when *Homo sapiens* play this game, they tend to attain the analytical PSE, although cooperation has been found to occur in some experiments (c.f., Heurer and Orland, 2019). This should come as no surprise. As we now know, the Prisoner's Dilemma (and *Homo sapiens* proclivity for attaining the game's PSE) is a contributing factor to historically intractable resource management problems in everyday life like air pollution, water scarcity, and climate change.[3]

FINITELY REPEATED PRISONER'S DILEMMA

Similar to the question of whether repeatedly playing the trust game for multiple stages could lead to greater trust and trustworthiness among an Investor and Trustee, the question arises as to whether repeated play of the Prisoner's Dilemma can lead to more cooperation among two players in a PSE. Unfortunately, applying backward induction to the payoff matrix above for a finite number of periods suggests that the answer is "no." To see this, suppose two players are in the final stage of the game. Given the payoff matrix above, (Deviate, Deviate) is the PSE. Knowing this, in the penultimate stage, both players have no better options than to choose (Deviate, Deviate) again. Similar to the centipede game where mutual mistrust unfolds all the way back to the initial round, here, mutual deviation unfolds back to the first stage. Analytically speaking, cooperation does not emerge in a finitely repeated Prisoner's Dilemma game—repeatedly deviating is the dominant strategy for both players.

2. Game theorists make a distinction between "strictly" dominant strategies and "weakly" dominant strategies. In the Prisoner's Dilemma, the strategies are strictly dominant because (1) the payoff of choosing Deviate when the other player Deviates (\$2) is greater than choosing Cooperate (\$1), and (2) the payoff of choosing Deviate when the other player chooses Cooperate (\$6) is also greater than choosing Cooperate (\$4). If either (not both) of these values were equal to each other (e.g., the \$2 payoffs in the (Deviate, Deviate) cell were instead equal to \$1 each, or the \$4 payoffs in the (Cooperate, Cooperate) cell were instead equal to \$6 each), then the Deviate strategy would only be weakly dominant.
3. Poundstone (1992) provides the onomatology of the title Prisoner's Dilemma. As the title suggests, Prisoner's Dilemma was used to describe a fictional game where two suspects are apprehended, and the investigator wants both to individually confess to having participated in a crime. The investigator sets the prison sentences associated with confessing (Deviate) and maintaining innocence (Cooperate) such that the suspects' dominant strategies are to confess.

Nevertheless, Kreps et al. (1982) and Andreoni and Miller (1993), among others, find that *Homo sapiens* tend to cooperate more when they are uncertain of the other player's tendency to cooperate. Furthermore, Spaniel (2011) demonstrates that when players adopt strategies such as "grim trigger" and "tit-for-tat" in an infinitely repeated Prisoner's Dilemma game, cooperation is more likely to occur. Grim trigger is a strategy where if your opponent Deviates at any stage, you Deviate forever starting in the next stage. Otherwise, you continuously Cooperate. Tit-for-tat is where you begin by choosing to Cooperate. In future stages, you then copy your opponent's play from the previous period.[4]

PUBLIC GOOD GAME

As in the Prisoner's Dilemma, the dominant strategy in a Public Good Game results in a PSE among *Homo economicus,* and often among *Homo sapiens*, that is inefficient when compared with what would otherwise be a cooperative outcome. In a simple version of this game—called a linear public good game—there is a group of N players, each of whom receives an initial allocation of money, say \$10, and then is asked how much of that \$10 she will voluntarily contribute to a group project of some kind. Each dollar that is donated by an individual player to the group project is multiplied by some factor α, $1 < \alpha < N$, and shared equally among all members of the group (for the sake of a game played in a laboratory, the group project is simply a pot of money). The fact that $1 < \alpha < N$ ensures that an individual player's contribution to the pot of money is larger in value for the group as a whole than it is for that individual.

For example, suppose there are four players (including you) and α = 2. If you contribute a dollar to the pot, then you give a dollar and receive only \$0.50 out of the pot in return ((\$1 x 2) ÷ 4 players = \$0.50). However, for every dollar contributed by another player, you receive \$0.50 out of the pot, free and clear. You can see how this game mimics the social dilemma *Homo sapiens* face when it comes to voluntarily financing a myriad of public goods (or group projects) such as public radio and public TV, environmental groups, and political campaigns, to name but a few. Each dollar contributed by someone else gives you additional benefit associated with a larger public good, for free. You get that same additional benefit from the public good when you are the one contributing but at the cost of your contribution.

Since the overall return the group gets from your dollar contribution exceeds the dollar (recall that in our case the overall return is \$1 x 2 = \$2 per dollar contributed), it is best (i.e., socially efficient) for each group member to contribute their full \$10 allocation. The socially efficient equilibrium to this game earns each player a total of \$20 ((\$10 x 2 x 4 players) ÷ 4 players = \$20). In contrast, because each player only gets back \$0.50 for each dollar contributed, there is no individual incentive for a player to donate any amount of money (or, to put it in economists' terms, there is a strong incentive for each player to "free ride" on the generosity of the other players' contributions). Thus, the PSE for this game—again, the equilibrium we expect *Homo economicus* players to obtain as a consequence of their self-interested, rational mindset—is where each player free rides and contributes nothing. Grim, but true.

One way to convince yourself that the PSE for the Public Good Game is where each player contributes zero to the group project is to start at some arbitrary non-zero contribution level for

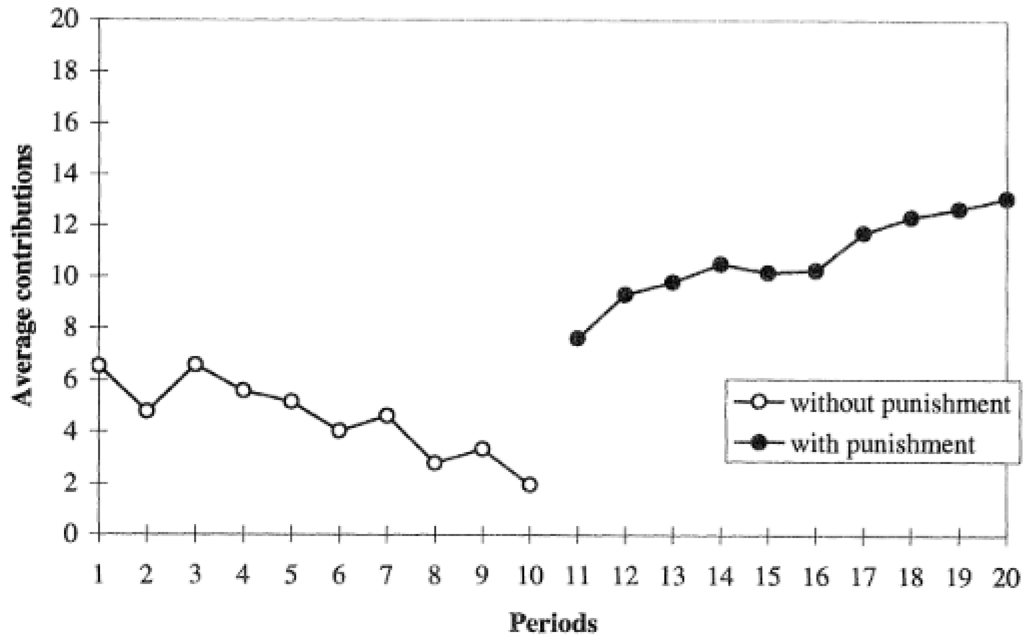
4. In a more recent experiment, where players engage in what they interpret as an infinitely repeated Prisoner's Dilemma game, Kim (2023) finds that higher discount factors (as a result of week- or month-long delays in when players are allowed to recoup their payoffs) induce more cooperation among the players.

each player and then show that each player has an incentive to reduce their contribution to zero. For example, suppose the starting point in our game is where each player contributes $6 to the money pot. This means that each player would receive a total of $12 from the pot (($6 x 2 x 4 players) ÷ 4 players =$12). Thus, each player takes home from the game a total of $12 + $4 = $16. Now suppose one of the players decides instead to free-ride by dropping her contribution to $5. Each of the four players now receives $11.50 from the money pot ([($6 x 2 x 3 players) + ($5 x 2 x 1 player)] ÷ 4 players = $11.50). The total take-home pay for the three non-free-riding players is now $11.50 + $4 = $15.50, while the free-riding player takes home $11.50 + $5 = $16.50. Clearly, the free-riding player is better off by having dropped his contribution to $5, and the three non-free-riding players are each worse off. But then each of the non-free-riders would recognize that they too would have been better off by free-riding, just like the free-rider. So, they too have equal incentive to free-ride. Barring any type of pre-commitment made by each of the players, this free-riding process cascades to each player choosing to fully free-ride or, to use the Prisoner's Dilemma lingo, to "deviate" from what is otherwise a fully cooperative equilibrium where each player contributes their total $10 to the money pot. *Voila*, we arrive at a PSE where each player chooses to contribute zero to the money pot.

Taking their cue from the likes of Kreps et al. (1982) and Andreoni and Miller (1993) in testing a finitely repeated version of the Prisoner's Dilemma Game, Fehr and Gächter (2000) explore whether a finitely repeated Public Good Game likewise mitigates deviation on the part of the players (i.e., free-riding behavior). The authors construct four treatment groups of student subjects. There is a "stranger treatment," with and without punishment opportunities, and a "partner treatment," with and without punishment (punishment in this context is explained below). In the partner treatment, 10 groups of four subjects each play the linear public good game for ten rounds without punishment and ten rounds with punishment, with group composition remaining unchanged across rounds (hence, the title, "partner treatment"). In contrast, in the stranger treatment, a total of 24 subjects are randomly partitioned into groups of four players in each of the twenty rounds (10 rounds without punishment and 10 rounds with punishment). Group composition in the stranger treatment changes randomly from round to round. In both treatments, subjects anonymously interact with each other.

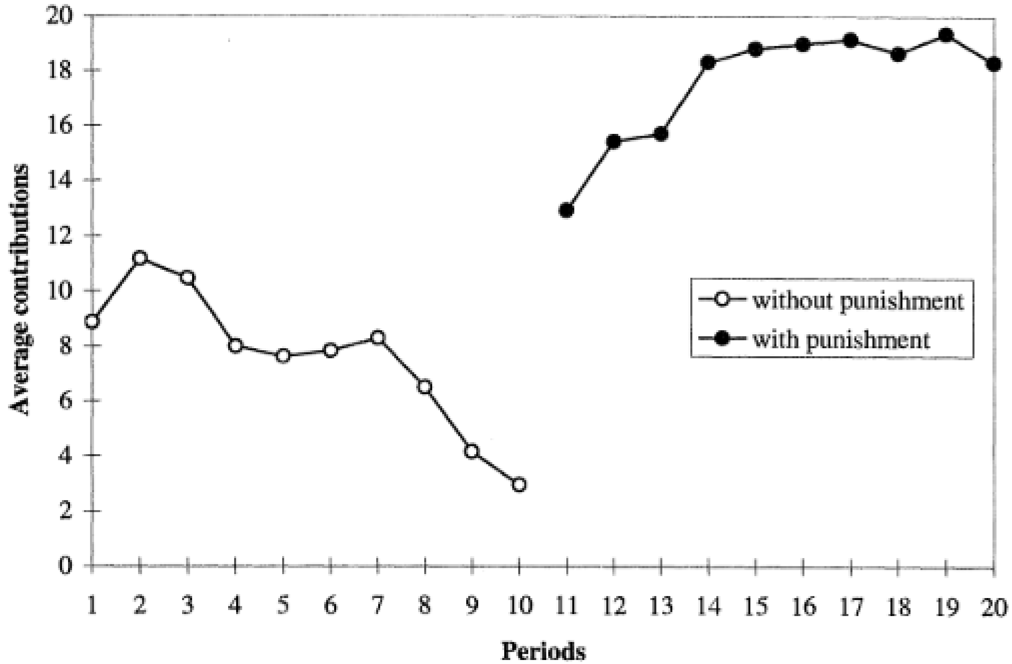
Games played without punishment opportunities serve as a control for games played with punishment opportunities. In a game with punishment opportunities, each subject is provided the opportunity to punish any other player (after any given round) after being informed about each player's contribution during that round. Punishment in this game takes the form of one player (the punisher) assigning "punishment points" to another player (the punished). For each punishment point assigned to a player, the player's payoff from that round is reduced by 10% (not to exceed 100%). To mitigate the potential misuse of the punishment mechanism, punishers face an increasing cost associated with assigning punishment points. The cost rises one-for-one with the first two points assigned, and then rises at an increasing rate for points assigned beyond two. Egads, this is sounding a bit complicated.

Fehr and Gächter's results are depicted in the following two figures: the first figure shows results for groups in the stranger treatment where the first 10 rounds are played without punishment, and the second 10 rounds are played with punishment. The expectation is that the availability of punishment opportunities would lead to an increase in the average player's contribution to the money pot. This is depicted in the first figure both by the discrete jump in contribution level starting in round 11, and the steady increase in this level over the remaining 10 rounds of the game. The downward trend in contribution levels over the first 10 rounds played indicates that the players learned that cooperation does not pay in a public good game without some form of punishment.



(Fehr and Gächter 2000)

The second figure shows similar results for the partner-treatment groups. What is notable in comparison with the results for the stranger-treatment groups is that (1) the downward trend in the initial 10 rounds becomes noticeably steeper from round seven onward, and (2) the initial jump up in average contribution level starting in round 11 (when punishment opportunities become available) is markedly larger, leading to higher per-round average contributions levels thereafter. These results demonstrate what is commonly known as "reputational effects" associated with a player's history of contributions over time. Among partner-treatment groups, reputational effects are enabled, while among stranger-treatment groups, they are not.



(Fehr and Gächter 2000)

It appears that punishment works with *Homo sapiens* in repeated play of a Public Goods Game, similar to how punishment works with *Homo sapiens* in repeated play of a Prisoner's Dilemma Game.[5]

In addition to enabling punishment opportunities as a "coordinating mechanism" to reverse the grim, inefficient, free-riding equilibrium among *Homo sapiens* in a finitely repeated Public Good Game without punishment, Rondeau et al. (1999) and Rose et al. (2002) propose a promising mechanism for one-shot games, called the Provision-Point Mechanism. As Rose et al. explain, a provision point mechanism solicits contributions for a public good by specifying a provision point, or threshold, and a money-back guarantee if total contributions do not meet the threshold. Extended benefits are provided when total contributions exceed the threshold. The authors report that the provision point mechanism has led to increased contribution levels (and thus adequate funding for public goods) in their laboratory and field experiments.[6]

5. In an intriguing cross-cultural comparison of the effectiveness of punishment in finitely repeated public good games, Gächter et al. (2010) find a surprising result regarding the punishment of above-average contributors. For example, in the Australian city of Melbourne and the European cities of Bonn, Nottingham, and Copenhagen, there is little punishment for above-average contributors, while in the Saudi Arabian city of Riyadh, the Greek city of Athens, and the Belarusian city of Minsk, above-average contributors were punished at the same levels as below-average contributors. The authors call this latter form of retribution anti-social punishment, attributed to their observation of a strong correlation between punishment received in one period and that doled out in the next. For example, if above-average Sam punished below-average Sally in period 1, Sally then punished Sam in period 2 as revenge. Because anti-social punishment is ultimately associated with less cooperation among players, this study's findings serve as a cautionary tale for those who espouse punishment as a universal remedy to the free-riding problem.
6. With their laboratory experiments, Chan et al. (2002) sought to answer the question of whether involuntary transfers for the provision of a public good, such as taxation, crowd out voluntary transfers (i.e., private donations) that reflect a "warm-glow" emotion among subjects. The authors do not find evidence of complete crowding out in general, but suggest that crowding out increases as the rate of taxation increases. Sufficiently large rates of taxation offset the benefits of warm-glow giving.

STAG HUNT

As its name suggests, this game tests the extent to which hunters can coordinate their efforts to bring down big game. Skyrms (2004) explains the game's onomatology—the Stag Hunt is a coordination game in which two hunters go out on a hunt together. Each can individually choose to hunt a stag or a rabbit. If one of the hunters hunts a stag, she must have the cooperation of the other hunter to succeed. Thus, like in the Prisoner's Dilemma, choosing to cooperate is risky—the other hunter can indicate he wants to cooperate but, in the end, take the less risky choice and go after a rabbit instead (remember, like in the Prisoner's Dilemma, the players' decisions are made simultaneously in this game). Alone, a hunter can successfully catch a rabbit, but a rabbit is worth less than a stag. We see why this game can be taken as a useful analogy for social cooperation, such as international agreements on climate change. An individual alone may wish to cooperate (e.g., reduce his environmental "footprint"), but he deems the risk that no one else will choose to cooperate as being too high to justify the change in behavior that his cooperation entails.

Here is the game's payoff matrix:

		Player 2	
		Stag	Rabbit
Player 1	Stag	3 , 3	0 , 2
	Rabbit	2 , 0	1 , 1

We use the same logic to determine this game's analytical PSE as we did to determine the Prisoner's Dilemma:

Player 1 first decides what to do if Player 2 chooses to hunt stag. Because Player 1's payoff in this case from hunting stag ($3) exceeds his payoff from hunting rabbit ($2), Player 1 will choose to hunt stag when Player 2 hunts stag. Next, we see that when Player 2 chooses to hunt rabbit, Player 1 will also choose to hunt rabbit since, in this case, the payoff from hunting rabbit ($1) exceeds the payoff from hunting stag ($0). Using the same approach to determine what Player 2's best strategy is, we see that she will also choose to hunt stag when Player 1 hunts stag and will hunt rabbit when Player 1 hunts rabbit. Hence, neither player has a dominant strategy in this game, and as a consequence, there are actually two PSEs. One PSE is where both players hunt stag; the other is where both hunt rabbit. We cannot say for sure which of the two equilibria will be obtained.

Clearly, the (stag, stag) equilibrium is preferable (also known as "Pareto dominant").[7] But this equilibrium requires that credible assurances be made by each player. In contrast, the (rabbit, rabbit) equilibrium is "risk dominant" in the sense that by choosing to hunt rabbit both players avoid the risk of having gone for stag alone. We would expect this equilibrium to occur when neither player is capable of making a credible assurance to hunt stag. Also note that even though this game does not permit the use of backward induction by the players (as a result of the game consisting of simultaneous rather than sequential moves), each player inherently uses "forward induction" in predicting what the other player will choose to do.

Belloc et al. (2019) recently conducted an experiment where a random sample of individuals playing a series of Stag Hunt games are forced to make their choices about whether to hunt stag or rabbit under time constraints, while another sample of players has no time limits to decide. The authors find that individuals under time pressure are more likely to play stag than individuals not under a time constraint. Specifically, when under time constraints, approximately 63% of players choose to hunt stag as opposed to 52% when no time limits are imposed.

ZERO-SUM GAME

Consider the following payoff matrix (where, as with the Prisoner's Dilemma and the Stag Hunt, each player's payoffs are common knowledge):

		Player 2	
		Left	Right
Player 1	Up	1 , -1	-1 , 1
	Down	-1 , 1	1 , -1

The reason why this matrix depicts a zero-sum game is because the payoffs to Players 1 and 2 sum to zero in each cell. Any time a player wins $1, the other player loses $1. You may have heard someone say, "my gain is your loss" or the other way around, or perhaps you've said something like this to

7. Vilfredo Pareto was a 19th – 20th-century Italian engineer, sociologist, political scientist, philosopher, and economist. He made several important contributions to economics, particularly concerning the study of income distribution and analysis of individuals' choices. He is considered one of the fathers of welfare economics.

someone yourself. When this happens, the two individuals are (unwittingly or not) acknowledging that they are participating in a zero-sum game.[8]

Using the same logic as we used in the Prisoner's Dilemma and Stag Hunt games to determine the player's best strategy, we find that there is no PSE for this game. Player 1's best strategy is to choose Up when Player 2 chooses Left, and Down when Player 2 chooses Right. On the contrary, Player 2's best strategy is to choose Right when Player 1 chooses Up, and Left when Player 2 chooses Down. No PSE emerges. What is *Homo economicus* to do?

It turns out that the analytical equilibrium solution concept for games such as this is what's known as a mixed-strategy equilibrium (MSE), where players choose probabilistic mixtures in which no single strategy is played all the time. For instance, if I always choose a particular strategy, and you anticipate that strategy, then you will win. I should, therefore, behave more unpredictably. Randomizing is a sensible strategy for me to follow when a little genuine unpredictability will deter the other player from making a choice that leads to a suboptimal outcome for me. The equilibrium involves unpredictable mixing on both the players' parts.

To facilitate the role randomization plays in determining an MSE, we amend the game's payoff matrix to account for each player's probabilistic moves.

		Player 2	
		q Left	$1-q$ Right
Player 1	p Up	1 , -1	-1 , 1
	$1-p$ Down	-1 , 1	1 , -1

Now we suppose that Player 1 chooses Up with probability p (and thus, Down with probability $(1-p)$) and Player 2 chooses Left with probability q (and Right with probability $(1-q)$). It turns out this game's MSE occurs when (1) Player 1 chooses p such that Player 2 is indifferent between choosing

8. As with the Prisoner's Dilemma and Stag Hunt games, environmental and resource economists recognize that the global fight against climate change exhibits features of a zero-sum game. All else equal, whenever one country invests in abating its carbon emissions, all other countries gain by not having to expend funds themselves to get the same amount of reduced carbon emissions. In terms of who bears the opportunity cost of the investment, this is one country's loss (or at least attenuated gain) and every other country's gain (since units of carbon reduced anywhere on the planet reduce the atmospheric stock of carbon that is responsible for rising ground temperatures and other meteorological changes occurring across the planet).

Left or Right (i.e., Player 2's expected payoff from choosing Left equals her expected payoff from choosing Right), and (2) Player 2 simultaneously chooses q such that Player 1 is indifferent between choosing Up or Down (i.e., Player 1's expected payoff from choosing Up equals his expected payoff from choosing Down). In particular:

Player 1 chooses p such that:

$(-1 \cdot p) + 1 \cdot (1 - p) = (1 \cdot p) + (-1 \cdot (1 - p)) \Longrightarrow p = \frac{1}{2}$.

Player 2 similarly chooses q such that:

$(1 \cdot q) + (-1 \cdot (1 - q)) = (-1 \cdot q) + 1 \cdot (1 - q) \Longrightarrow q = \frac{1}{2}$.

Thus, Player 1 (as a member of *Homo economicus*) chooses Up half the time, and Player 2 (also a member of *Homo economicus*) chooses Left half the time. Probably the best way each player can be true to their respective strategies is to flip a fair coin (e.g., for Player 1, it might be "Heads I go Up, tails I go Down," and similarly for Player 2).

Let's see why the equality $(-1 \cdot p) + 1 \cdot (1 - p) = (1 \cdot p) + (-1 \cdot (1 - p))$ holds when Player 2's expected payoff from choosing Left equals her expected payoff from choosing Right (you'll then be able to see why $(1 \cdot q) + (-1 \cdot (1 - q)) = (-1 \cdot q) + 1 \cdot (1 - q)$ holds when Player 1's expected payoffs from choosing Up and Down are equated). When Player 2 chooses Left and Player 1 chooses Up, Player 2's payoff is -\$1. The probability of Player 1 choosing Up is p, hence Player 2's expected payoff from choosing Left, conditional on Player 1 choosing Up, is $\$(-1 \cdot p)$. Similarly, when Player 2 chooses Left and Player 1 chooses Down, Player 2's payoff is \$1. The probability of Player 1 choosing Down is $(1 - p)$, hence Player 2's expected payoff from choosing Left, conditional on Player 1 choosing Down, is $\$(1 \cdot (1 - p))$. We then sum these two values (i.e., $\$(-1 \cdot p)$ plus $\$(1 \cdot (1 - p))$) to attain Player 2's (unconditional) expected payoff from choosing Left. The same process is followed to determine Player 2's expected payoff from choosing Right. Setting these two expected payoffs equal solves for $p = 1/2$.[9]

The proof for why an MSE is determined by each player randomizing their choice such that the expected payoffs for the other player are equated across that player's choices is simply proved by contradiction. If, for example, Player 1 randomizes his choices such that Player 2's expected payoff is larger when she chooses Left than Right (i.e., because Player 2 can see that $p < 1/2$), Player 2 will always choose Left. Because he more often chooses Down when $p < 1/2$, Player 1's payoff is lower than it otherwise would be if he instead chose p = 1.[10] The same logic holds when Player 1 chooses $p > 1/2$, in which case Player 2 always chooses Right. Because he more often chooses Up when $p > 1/2$, Player 1's payoff is again lower than it otherwise would be if he instead chose p = 1. This is because Player 1's payoff is again guaranteed to be -\$1 when $p > 1/2$. Thus, the best Player 1 can do is set p = 1. The proof is the same for Player 2. Hence, we have proved why an MSE is obtained—and is the optimal outcome for this game—when each player randomizes their choice such that the expected payoffs for the other player are equated across that player's choices. Whew!

Camerer (2003) informs us that in studies with two-choice games (i.e., games where the two players play the zero-sum game two times consecutively), *Homo sapiens* tend to use the same strategy after a win but switch strategies after a loss. This "win-stay, lose-shift" heuristic is a coarse version of what's known as "reinforcement learning." In four-choice games (players play four times consecutively),

9. To be clear, the MSE for the zero-sum game does not always result in $p = q = 1/2$. To see this, calculate the equilibrium values for p and q when the payoff matrix is something like (Up, Left) = (3,-3), (Up, Right) = (-2,2), (Down, Left) = (-1,1), and (Down, Right) = (0,0).

10. This is because Player 1's payoff is guaranteed to be -\$1 when $p < 1/2$.

Homo sapiens' strategies are remarkably close to *MSE* predictions by the fourth time they play the game.

When I was teaching in Myanmar, this was the one game I was called into service to play myself (we had an odd number of students that day). By then, I had gotten to know my students quite well individually. I was designated Player 1, and my student (I will call her Sally) was designated Player 2. We were playing a "single-choice game" due to time constraints imposed on the course. I had observed Sally over the previous weeks and concluded that she was unlikely to just flip a proverbial coin in deciding whether to choose Left or Right. She was right-handed and always sat to my right in the classroom. I guessed she would choose Right. So, I chose Down. My guess was, luckily for me, proven correct by Sally.

Afterward, when I explained my strategy to the students, I emphasized that a player need not actually randomize his moves as long as his opponent cannot guess what he will do. An MSE can therefore be an "equilibrium in beliefs," beliefs about the likely frequency with which an opponent will choose different strategies. But I reminded the students that we had only played a single-choice game. With repeated play, chances are Sally would have begun to randomize her choices, to the point that flipping a coin would become my best strategy as well. We would slowly but surely evolve from *Homo sapiens* to *Homo economicus*.

STAG HUNT (REPRISE)

In our first assessment of the Stag Hunt game, we learned that two PSEs exist, with no way of definitively determining which of the two are most likely to occur. As a result, we are compelled to determine the game's MSE, as this is as close as we can get to identifying a unique analytical equilibrium. The game's payoff matrix is reproduced here, this time accounting for the players' probabilistic moves:

		Player 2	
		q	$(1-q)$
		Stag	Rabbit
p	Stag	3 , 3	0 , 2
Player 1			
$(1-p)$	Rabbit	2 , 0	1 , 1

Using the same procedure as shown for determining the analytical MSE for the Zero-Sum game, the value of p for Player 1 is determined as $3p = 2p + (1-p) \Longrightarrow p = 1/2$ and the value of q for Player 2 is determined as $3q = 2q + (1-q) \Longrightarrow q = 1/2$. Each hunter might as well flip a fair coin in deciding whether to hunt stag or rabbit.

It turns out that the MSE for this game results in expected payoffs for each player that are larger than the certain payoffs obtained when both hunt rabbit (which, you'll recall, is one of the game's PSEs), but lower than the payoffs obtained when both hunt stag (the game's other PSE). To see this, we can calculate Player 1's expected payoff for the MSE as $\left(3 \cdot \frac{1}{4}\right) + \left(0 \cdot \frac{1}{4}\right) + \left(2 \cdot \frac{1}{4}\right) + \left(1 \cdot \frac{1}{4}\right)$ = \$1.5. Dissecting this equation, the first term is Player 1's payoff in the (Stag, Stag) cell of the matrix multiplied by the probability that both players will choose to hunt stag ($p \cdot q = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$). Similarly, the second term—Player 1's payoff in the (Stag, Rabbit) cell multiplied by the probability that Player 1 chooses to hunt stag and Player 2 chooses to hunt rabbit—is equal to $0 \cdot p \cdot (1-q) = 0 \cdot \frac{1}{2} \cdot \frac{1}{2} = 0 \cdot \frac{1}{4}$. And so on for Player 2. This is another way of saying that, when faced with the Stag Hunt, flipping a coin essentially leads to a bit less than "splitting the difference" for each player from jointly hunting stag and jointly hunting rabbit (technically speaking, splitting the difference would result in payoffs of \$2 each).

Cooper et al. (1990 and 1994) used the following payoff matrix as the baseline (or what they call the Control Game (CG)) for their Stag Hunt game experiment:

		Player 2	
		Stag	Rabbit
Player 1	Stag	1,000 , 1,000	0 , 800
	Rabbit	800 , 0	800 , 800

A clear majority of their *Homo sapiens* pairs who participated in this single-choice CG game obtained the inefficient (Rabbit, Rabbit) equilibrium. The authors also had different sub-groups of subjects play what they called (1) the CG-900, where Player 1 could opt out and award both players 900 instead of playing the game (note that 900 > 800); (2) the CG-700, where Player 1 could opt out and award both players 700 instead of playing the game (note that 700 < 800); (3) CG-1W, where one of the two players is allowed to engage in "cheap talk" with the other player, presumably to nudge the other player into committing to hunt stag; and (4) CG-2W, where both players are allowed to engage in cheap talk in an effort to nudge each other into committing to hunt stag.

Cooper et al. found that 97% of player pairs chose the (Rabbit, Rabbit) PSE in the CG treatment. A large number of Player 1s also took the outside option in the CG-900 treatment. In cases where Player 1 refused the outside option, more than a supermajority of player pairs (77%) obtained the efficient (Stag, Stag) equilibrium. In the CG-700 treatment, the majority of player pairs reverted to the inefficient (Rabbit, Rabbit) equilibrium. Lastly, with one-way cheap talk between players the efficient (Stag, Stag) equilibrium jumps from 0% in the CG treatment to 53%. Unexpectedly, the jump is even greater for two-way cheap talk (up to 91%).

These results are encouraging for *Homo sapiens* because by simply allowing players to communicate with each other (presumably pre-committing to hunt stag), *Homo sapiens* are, for the most part, capable of attaining the efficient outcome where both players hunt stag together. In cases where an outside option is available for one of the players, as long as that option's payoff is larger than the mutual payoffs associated with the (Rabbit, Rabbit) PSE, yet lower than the mutual payoffs associated with the (Stag, Stag) PSE, the player with the outside option conjectures that the other player will choose to hunt stag, who is likely to end up confirming that conjecture.

BATTLE OF THE SEXES

While it is unclear who actually named this game, there is little debate about the genesis of the title's popularization, which occurred on Mother's Day in 1973 at the dawn of the women's liberation movement.[11] Tennis stars Bobby Riggs and Margaret Court faced off in a $10,000 winner-take-all challenge match, which then 55-year-old Riggs, a tennis champion from the late 1930s and 40s who was notoriously dismissive of women's talents on the tennis court, resoundingly won. Later that year, Riggs challenged the higher-profile tennis star Billie Jean King to a $100,000 winner-take-all challenge match, which King won handily.

Although the game we have in mind here is far from being a sports match between the sexes, it does capture the flavor of challenges that sometimes bedevil couples' coordination decisions. The payoff matrix for this game is presented below. Like the Stag Hunt, the game's analytical equilibrium consists of two PSEs (can you identify them?). Therefore, to determine the game's unique MSE, we acknowledge Spouse 1's and Spouse 2's probabilistic strategies upfront.

11. The following interpretation is taken from history.com (2020).

		Spouse 2	
		q Ballet	$(1-q)$ Martial Arts
Spouse 1	p Ballet	1 , 2	0 , 0
	$(1-p)$ Martial Arts	0 , 0	2 , 1

Given what you have just learned about solving for MSEs in the Stag Hunt and Zero-Sum games, you can show that for this game, p = 1/3 and q = 2/3. Further, the expected payoff in the MSE for each spouse is \$0.67. Interestingly, these expected payoffs are lower than the least preferable payoff in either of the game's two PSEs, where the spouses either agree to watch the martial arts performance or attend the ballet. Recall that in the Stag Game's MSE the expected payoffs for each player "split the difference" of payoffs from the two PSEs.

In Cooper et al.'s (1989) experiments with *Homo sapiens*, subjects mismatched 59% percent of the time, which is actually an improvement over mismatching in the game's analytical MSE. Mismatching occurs in the MSE when one player chooses Ballet and the other Martial Arts. This occurs $p\,(1-q)+q\,(1-p)=\frac{1}{9}+\frac{5}{9}$ = 67% of the time. The authors also found that when one player (say, Spouse 1 in our game) is given an *outside option* (which, if Spouse 1 takes, pays him and his spouse some value x, such that $1<x<2$), and the husband rejects the option, the analytical equilibrium would entail Spouse 2 surmising that Spouse 1 will then choose Martial Arts. Thus, Spouse 2 should also choose Martial Arts. In their experiment, Cooper et al. (1989) found that only 20% of Spouse 1s chose the outside option. Of the 80% of Spouse 1s who rejected the option, 90% obtained their preferred outcome—Martial Arts, here we come!

With one-way communication, the players coordinated their choices 96% of the time! However, with simultaneous two-way communication, they coordinated only 42% of the time! What happened? Recall that, in the Stag Hunt game, two-way communication enhanced coordination. Here, in the Battle of the Sexes, it has the opposite effect. Lastly, Cooper et al. (1989) found that when one of the players is known to have chosen ahead of time, but the other player is not informed about what the other player chose, the mismatch rate between the players decreased by roughly half (relative to the baseline game with no communication or outside options).

PENALTY KICK

There are few team sports where an individual player is put in as precarious a position as a goalie defending a penalty kick in the game of soccer (or football if you are from anywhere else in the world except the US, Canada, Australia, New Zealand, Japan, Ireland, and South Africa). *Homo sapiens* rarely look more vulnerable than when put in a position of having to defend a relatively wide and high net from a fast-moving ball kicked from only 12 yards away.

Spaniel (2011) provides a nice analogy in the context of a payoff matrix where we are forgiven for taking the liberty of depicting the analytical equilibrium as an MSE upfront. Do the payoff combinations for the striker and goalie in each cell of the matrix ring a bell? The bell should be ringing zero-sum game.

		Goalie q L	Goalie $(1-q)$ R
Striker	p L	$0, 0$	$x, -x$
Striker	$(1-p)$ R	$1, -1$	$0, 0$

For this game, we assume a superhuman (as opposed to a mere *Homo economicus*) goalie. If the striker kicks Left (L) and the goalie guesses correctly and dives L, the goalie makes the save for certain. Similarly, if the striker kicks Right (R) and the goalie correctly dives R, the goalie again makes the save for certain. The striker, however, is fallible. If she kicks R and the goalie dives L, she scores for certain. But if she kicks L and the goalie dives R, she only scores with probability x.

Using the method we previously developed to solve for an MSE in the Stag Hunt, Zero-Sum, and Battle of The Sexes games, verify that, in the equilibrium, $p = \frac{1}{1+x}$ and $q = \frac{x}{1+x}$. Further, for those of you who know a little calculus, you can use these equations to solve for the respective first partial derivatives of p and q with respect to x, as $\frac{\partial p}{\partial x} = -\frac{1}{(1+x)^2} < 0$ and $\frac{\partial q}{\partial x} = \frac{1}{(1+x)^2} > 0$. Herein lies the closest thing to understanding the likely MSE choices that are made by the goalie and striker.[12]

12. Note that solving for the likelihood that a goal will actually be scored on any given penalty kick is, to put it mildly, anyone's guess.

The first partial derivatives together inform us that, in an MSE, as the striker's probability of scoring goes up when she kicks L (embodied by an increase in x), the probability of the striker kicking L actually goes down. Analytically speaking, it is as if the striker uses one degree of iterated knowledge to determine the kick's direction. The more the goalie believes the striker has a higher probability of scoring when she kicks L, the more likely the goalie will dive L. Thus, it makes more sense for the striker to kick R.

The next time you watch a game with lots of penalty kicks, you will be able to test this equilibrium concept in your own experiment with *Homo sapiens*.

HOTELLING'S GAME

No, this is not a game played among hoteliers. The game is named after Harold Hotelling, a 20th-century mathematical statistician and theoretical economist who pioneered the field of spatial, or urban, economics. The game is depicted below:

Suppose there are two vendors (**v1** and **v2**) on a long stretch of beach selling the same types of fruit juices. The vendors simultaneously choose where to set up their carts each day. Beachgoers are symmetrically distributed along the beach. The beachgoers buy their fruit juice from the nearest vendor.

The line graph below distinguishes the furthest south location on the beach at 0 and the furthest north location at 1.

If you are one of the two vendors where would you decide to locate your cart on the beach?

The analytical equilibrium for this game is a PSE. The logic behind its solution goes like this:

> *If v1 locates at ½ she guarantees herself at least half of the total amount of business on any given day—more if v2 doesn't also locate at ½. Indeed, v1 locating at ½ maximizes her profit. To see this, start v1 at 0 and v2 at 1. Hold v2 at 1 and move v1 toward ½. Note that v1 commands the most exposure to beachgoers at location ½. Given that v1 chooses location ½, it is in v2's best interest to also locate at ½ (using the same logic we used to determine that v1 would locate at ½). Thus, (½, ½) is the PSE for this game.*

It should be no surprise that this result is also known as the Median Voter Theory: throughout a typical campaign, candidates for public office tend to gravitate toward the middle of the political spectrum, i.e., toward the "median voter."

Collins and Sherstyuk (2000) studied how two-, three-, and four-player Hotelling Games were played among *Homo sapiens*. The authors found that, in two-player games, *Homo sapiens*' strategies are remarkably close to the analytical PSE prediction of (½, ½). In four-player games, the analytical PSE occurs at locations ¼ and ¾. Homo sapiens cluster at these locations, but also somewhat in the middle as well. In three-player games, the PSE occurs where each player randomizes his locational choice uniformly over the interval of locations between ¼ and ¾ inclusive, denoted [¼, ¾]. Thus, location intervals [0, ¼) and (¾, 1] are avoided. In three-player games, relatively smaller percentages of players locate outside the PSE interval of [¼, ¾] and larger percentages of players locate within the PSE interval—way to go *Homo sapiens*! *Homo economicus* would have located strictly within the PSE interval, with each player randomizing his choice uniformly over the PSE interval.

IS MORE INFORMATION ALWAYS BETTER?

The rational-choice model of *Homo economicus* answers "yes", in most cases. The more information the better—more information can aid a consumer's decision-making. But wait. Spaniel (2011) offers a game where the answer is surprisingly "no," more information is not always better, even for *Homo economicus.*

Consider the following game: Player 1 chooses whether to play or not. If Player 1 chooses not to play, both Players 1 and 2 get $100. Simultaneously to Player 1's choice, Player 2 chooses between Heads (H) and Tails (T) on a coin flip, or chooses not to gamble on the coin flip. If Player 1 had chosen to play and Player 2 had chosen not to gamble on the coin flip, both players receive $200. If instead Player 2 chooses to gamble, the coin is flipped. If Player 2 has called the outcome of the flip correctly, she wins $300 and Player 1 loses $300; vice-versa if Player 2 has called the outcome of the flip incorrectly. Therefore, the game looks like this:

		Player 2		
		H	T	Quit
Player 1	Play	0, 0	0, 0	200, 200
	Quit	100, 100	100, 100	100, 100

Note that the (0,0) payoffs in the cells (Play, H) and (Play, T) represent the players' respective expected payoffs since whether the players win or lose $300 depends upon Player 2's 50-50 luck in correctly predicting the outcome of the coin flip. The analytical equilibrium for this game goes as follows:

> *Because Player 2's weakly dominant strategy is to Quit, the PSE for this game is for Player 1 to choose Play and Player 2 to choose Quit (i.e., (Play, Quit)). Note that this equilibrium is efficient! Most importantly for our purposes, Player 2 gets $200 in this equilibrium.*

Now we consider a slight tweak to this game where Player 2 has private information about the outcome of the coin flip before he decides whether to play or quit. In other words, Player 2 now knows whether the coin has come up Heads or Tails beforehand. Thus, Player 2 is assured of winning $300 if Player 1—who has no prior knowledge of the outcome of the coin flip—decides to Play. The key question is whether Player 1 will now choose Play with positive probability (i.e., $p > 0$). If not, then Player 2 has been harmed by having private information about the outcome of the coin flip—he wins only $100, instead of $200.

Recall that, in the previous game, Player 2 chose whether to play or quit before the coin was flipped. She was, therefore, uninformed about the outcome of the coin flip before deciding whether to play the game. Now, suppose Player 2 chooses whether to play or quit after the coin is flipped, and that the outcome of the coin flip is Player 2's "private information." The players are now involved in what's known as a "Bayesian Nash Game," where there are effectively two types of Player 2's—an H type (with a probability of q = 0.5 and a T type (with a probability of (1 – q) = 0.5. The game now looks like:

H Type ($q = \frac{1}{2}$)

Player 1		H	T	Quit
	Play	0 , 0	0 , 0	200, 200
	Quit	100, 100	100, 100	100, 100

H Type ($(1 - q) = \frac{1}{2}$)

Player 1		H	T	Quit
	Play	0 , 0	0 , 0	200, 200
	Quit	100, 100	100, 100	100, 100

The analytical equilibrium goes as follows:

Unfortunately for Player 2, Player 1 will set p = 0. If Player 1 sets $p > 0$, then Player 2 will simply play the correct side of the coin—H if it was H and T if it was T—thus ensuring a win of $300 when Player 1 Plays and winning $100 when Player 1 Quits. But then, Player 1 "wins" -$300 when she Plays, implying that Player 1 will never choose to Play (i.e., she will never choose $p > 0$). Thus, the analytical equilibrium for this game is (Quit, Quit) with each player winning $100.

Since $100 is less than what Player 2 won when he did not possess private information about the outcome of the coin flip (which was $200), more information in this context is not better for either player. I don't know about you, but there are plenty of instances where having less information to sift through before making a decision eases my mind and actually makes me feel happier. For example, I tend to assemble appliances, furniture, or equipment with much less angst when the instructions are concise and to the point, preferably accompanied by clear pictures or diagrams for me to follow. Lengthy written descriptions often prey upon my insecurities and impatience.

MARKET ENTRY

Consider the following game proposed by Camerer (2003):

Each of 20 players decides privately and anonymously whether to enter or stay out of the market. For each period, a different "market carrying capacity," denoted by odd integer $c \in$ (1, 3, 5, 7, 9, 11, 13, 15, 17, 19), is publically announced, after which the players make their entry decisions into the market.

Each player's payoff (v) is calculated as:

$$v = \begin{cases} k & \text{if the player stays out of the market} \\ k + r(c - m) & \text{if the player enters the market} \end{cases}$$

where k = \$100, r = \$2, and m represents the total number of players who enter the market. Note that at the time of their decisions, each player knows the values of k, r, c, and n = 20, but obviously not m.

Example: If c = 7, and in equilibrium m = 4, a player who decided not to enter the market that period receives \$100, and a player who entered the market receives 100 + (2 x 3) = \$106.

As Camerer (2003) shows, the analytical equilibrium for this game is rather complicated.

At the market level, there are two Nash Equilibria per value of c:

1. m = c*. For example, let* c *= 7. If* m *= 7, then each of the 20 players earns \$100. If one of the seven entrants instead chooses to stay out of the market, then she will not increase her payoff above \$100 (she now earns \$100 as a non-entrant). If, instead, one of the game's 13 non-entrants decides to enter the market, this new entrant will decrease her payoff from \$100 to \$98. Thus, no player has a "profitable deviation" from this equilibrium.*
2. m = $c - 1$*. For example, let* c *= 9 . If* m *= 8, then each of the eight entrants earns \$102, and each of the 12 non-entrants earns \$100. If one of the eight entrants instead chooses to stay out of the market, she will decrease her payoff by \$2 (from \$102 to \$100). If one of the game's 12 non-entrants decides to enter the market, then this new entrant's payoff remains \$100 (since now* c *=* m *= 9). Thus, again no player has a profitable deviation from the equilibrium.*

At the individual player level, there is a unique MSE. Letting $p\left(c\right)$ *represent the probability of any given player deciding to enter the market for a given* c*, and letting* n *represent the total number of players, a given player's expected payoff from entering is calculated using the expression for the binomial distribution,*

$$\sum_{s=0}^{n-1} \binom{n-1}{s} p\left(c\right)^{s} \left(1 - p\left(c\right)\right)^{n-1-s} \left\{k + r\left(c - s - 1\right)\right\}$$

For background on the binomial distribution see, [*http://onlinestatbook.com/2/probability/binomial.html*](http://onlinestatbook.com/2/probability/binomial.html).

Thus, an individual player reaches the MSE when,

$$\sum_{s=0}^{n-1} \binom{n-1}{s} p\left(c\right)^{s} \left(1 - p\left(c\right)\right)^{n-1-s} \left\{k + r\left(c - s - 1\right)\right\} = k$$

(i.e., when the player is indifferent between entering the market (expected payoff represented by the left-hand side of the equality) and staying out of the market (certain payoff represented by the right-hand side of the equality), which can be solved for

$$p(c) = \frac{[r(c-1)+k-v]}{r(n-1)}.$$

Thus, in equilibrium,

$$m = np(c) = \frac{n[r(c-1)+k-v]}{r(n-1)}.$$

In their laboratory experiments, Sundali et al. (1995) found that *Homo sapiens* mimic the analytical equilibrium rather closely. In a baseline setting (Experiment 1), 20 subjects were provided with no feedback between 60 rounds of play—a series of six blocks, each block based upon a randomly chosen value of $c \in (1, 3, 5, 7, 9, 11, 13, 15, 17, 19)$, were played by each subject 10 times each.[13] In the first block, one of the 20 subjects chose to enter the market (m = 1) when c = 1, four subjects entered (m = 4) when c = 3, seven entered (m = 7) when c = 5, and so on.

Sundali et al. ultimately find a relatively close correspondence between the mean values of m and their associated values of c across the six blocks. Recall that in the analytical equilibrium $m \approx c$ (i.e., m and c should be roughly equal). This close correspondence is supported by relatively large (close to one) correlation coefficients reported for each block (with a coefficient of 0.92 for the 60 rounds in total). This means that on average m and c moved in roughly the same direction over the 60 rounds—when c was a relatively large value, so was m, and when c was relatively small, so was m (0.92 $\approx$ 1, where 1 implies perfect linear correlation between the two values). Second, variability in the market entry decision across participants was largest for intermediate levels of c.

In a second experiment (Experiment 2), subjects were provided with feedback at the end of each round regarding the equilibrium value of m, as well as their respective payoffs for each round and cumulative payoffs up to the round. Subjects were also encouraged to write down notes concerning their decisions and outcomes. Results for this experiment were even closer to those predicted for *Homo economicus* in the analytical equilibrium.

As Camerer (2003) points out, how firms coordinate their entry decisions into different markets is important for an economy. If there is too little entry, prices are too high and consumers suffer; if there is too much entry, some firms lose money and waste resources, particularly if fixed costs are unsalvageable. Public announcements of planned entry could, in principle, coordinate the right amount of entry, but announcements may not be credible because firms that may choose to enter have an incentive to announce that they surely will do so in order to ward off competition. Government planning may help reduce this perverse incentive, but governments are nevertheless vulnerable to regulatory capture by prospective entrants seeking to limit competition. Evidence from the real world often suggests that too many firms choose to enter markets in general, particularly in newly forming markets. As we learned in Section 1, the phenomenon of excessive entry could reflect a Confirmation Bias on the part of potential entrants, leading to overconfidence in their abilities to obtain positive profits.[14]

WEAKEST LINK

Consider the following game presented in Camerer (2003):

13. Thus, each subject was presented with each value of c six different times over the course of the experiment.

14. Camerer and Lovallo (1999) designed an experiment to detect the extent to which overconfidence in one's skills relative to what he perceives are the skills of potential competitors induces excess entry into a market and subsequent business failure. The authors uncover a phenomenon they termed reference group neglect, where excess entry is much larger when subjects participated in the experiment knowing that payoffs would depend upon skill level. These self-selected subjects neglected the fact that they are competing with a reference group of subjects who all think they are as skilled.

Players (more than two) pick numbers from 1 to 7. Payoffs increase with the minimum of all the respective numbers chosen, and decrease with the deviation of their own choice from the minimum. The specific payoffs (in dollars) for our game are shown in the table below:

Lowest Choice in Group

Your Choice of *X*	7	6	5	4	3	2	1
7	13	11	9	7	5	3	1
6	---	12	10	8	6	4	2
5	---	---	11	9	7	5	3
4	---	---	---	10	8	6	4
3	---	---	---	---	9	7	5
2	---	---	---	---	---	8	6
1	---	---	---	---	---	---	7

(Camerer 2003)

For example, if a player chooses 4 and the minimum chosen by another player in the group is 2, then that player's payoff is $6.

You may recognize this game as being a generalization of the Stag Hunt and reminiscent of the Continental Divide. As such, there is an efficient equilibrium where each player chooses the number 7, and an inefficient, risk-dominant equilibrium where each player plays it safe by choosing the number 1. Choices of 5 and 6 act as potential basins of attraction for 7, while choices of 2 and 3 are potential basins of attraction for 1.

Van Huyck et al. (1990) played this game with a total of 107 *Homo sapiens* parcelled into groups consisting of between 14 and 16 subjects for 10 rounds. The authors found that relatively large numbers of the subjects choose higher numbers in the first round—33 players chose the number seven, 10 players chose the number six, 34 chose five, and 17 chose 4. However, by the tenth and final round, 77 players chose the number one and 17 chose the number 2. These results are obviously discouraging. Recall that the efficient equilibrium is where there is no weak link among the players—every player chooses the number seven, and no players choose the number one.

Interestingly, results for 24 *Homo sapiens* who played the game for only seven rounds were more encouraging. In the first round only 9 players chose number seven, zero chose number 6, and four chose number 4. By the seventh and final round, 21 had chosen the number seven.

Perhaps these results are unsurprising given what we've already learned about *Homo sapiens'* behavior in the Finitely Repeated Trust (Centipede) Game. In that game, we learned that in repeated play, trust and trustworthiness between two players can be achieved, at least temporarily. Random re-matching obviates the build-up of trust and trustworthiness that can arise through repeated play. And, generally speaking, it is more difficult to build trust and trustworthiness among a larger group of players.

As for real-world applications of the Weakest Link Game, Camerer (2003) points out that in the airline business, for example, a weakest link game is played every time workers prepare an airplane for departure. The plane cannot depart on time until it has been fueled, catered, checked for safety, passengers have boarded, and so on. For short-haul carriers, which may use a single aircraft on multiple flights daily, each departure is also a link in the chain of multiple flights, which creates another weakest link game among different ground crews.[15]

WEAKEST LINK WITH LOCAL INTERACTION (TWO VERSIONS)

We now consider two different versions of the Weakest Link Game presented in Camerer (2003). In the first version, players are provided with some information at the end of each round about the choices made by the other players. In the second version, players play separate games simultaneously with their neighbors, thus permitting the experimenter to test whether spatiality might affect the game's equilibrium outcome among *Homo sapiens*.

Here is version one:

> To begin, players individually choose the letter X or Y in separate groups of three players each over 20 rounds. Payoffs (in dollars) are shown in the table below. Each player learns the two choices made by the other two players at the end of each round (but not which player made with choice).

15. This suggests that efficient equilibria have been obtained in an airline company's weakest link games if all of a passenger's connecting flights are on time. Whether or not any given passenger makes it to the airport and his connecting flights on time is his set of weakest link games.

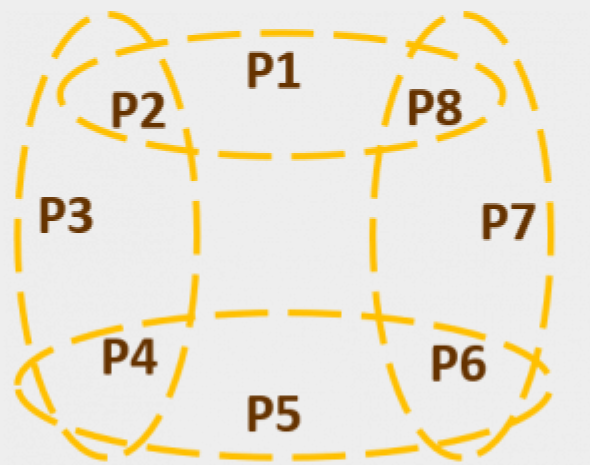
		Two Other Players		
		Both X	One X, One Y	Both Y
Row Player	X	80	60	60
	Y	10	10	90

And here is version two:

In this version of the game, players face the same payoff matrix:

		Two Other Players		
		Both X	One X, One Y	Both Y
Row Player	X	80	60	60
	Y	10	10	90

But each group of eight players is arranged in a circle, and each subject plays with his two other nearest neighbors:



For example, in this diagram Player P1 plays with Players P2 and P8 directly, while Player 2 simultaneously plays with Players P3 and P4 and Player P8 simultaneously plays with Player P6 and P7.

Note that the risk-dominant, inefficient equilibrium in both versions occurs when each player decides to play it safe by choosing option X. The efficient equilibrium occurs when all three players in a group choose Y. In this way, the game mimics the Weakest Link Game introduced earlier. In keeping with Van Huyck et al.'s (1990) original findings, we would therefore expect that, because the game here is being played in relatively small groups (three players per group), *Homo sapiens* playing these games would again mimic *Homo economicus* and obtain the efficient equilibrium, especially in version one of the game. Version two presents a complication since, although the groups are small, each player now plays with multiple sets of different players simultaneously, either directly or indirectly. The even-numbered players do so directly (e.g., Player P2 simultaneously plays with Players P1 and P8, on the one hand, and Players P3 and P4 on the other). Odd-numbered players do so indirectly (e.g., Player P1 plays with Players P2 and P8, who in turn play simultaneously with Players P3 and P4 and Players P6 and P7, respectively).

In playing this game with different groups of *Homo sapiens*, Berninghaus, et al. (2002) found that in standard (non-circular) three-person groups, players initially play Y about three-quarters of the time—specifically, in their experiments seven out of eight groups of three players coordinated on the efficient (all-Y) equilibrium. No surprise. However, in the local-interaction (circular) groups, players started by playing Y only half of the time, but then play of Y fell steadily to almost none by the 20th round of the game. Specifically, players responded to their neighbors by playing X 64% of the time when one other neighbor had chosen X. The authors likened this result to the spread of a disease through a population by close contact—in this case, the spread of fear. The incidence of playing X spreads from neighbor to neighbor, eventually infecting the entire group.

A merger where one or more disparate groups of individuals merge into a single group is an extreme form of local interaction, one that is continually played out through history on a macro-scale (e.g., beginning with families merging into clans who merge into ethnic groups, to villages and city-states to nation-states and sometimes empires). On a micro- scale, mergers occur when one company acquires another and becomes its "parent company." Mergers and acquisitions among businesses are a common feature of market economies—some would say they pose a threat to competition, while others believe they can provide efficiency gains via economies of scale.

Camerer and Knez (1994) wondered how the Weakest Link game might inform us about the efficacy of mergers occurring between two small groups into a single larger group. The group-size effect noted earlier suggests that, all else equal, larger (merged) groups would be less likely to converge to the efficient equilibrium where all group members individually select the number seven. This is exactly what Camerer and Knez found. Interestingly, the provision of public information (to each group about the other group with which it later merged) appears to have worsened the outcomes (i.e., increased the inefficiency) in paired small groups of three players each.

Camerer and Knez conclude that the results for merged groups in the context of a Weakest Link game are not encouraging. However, not to be outdone by these unflattering results, the authors added another treatment to the mix (in addition to the provision of public information about the performance of the other groups in the previous round). The additional treatment was a public announcement to all groups of a "bonus" if everyone picked the number 7 (i.e., if each merged group chose the Pareto efficient outcome). Thankfully, the announcement nudged the merged groups immediately in round 1, from 90% choosing numbers one or two to 90% choosing number 7. Camerer and Knez's *Homo sapiens* apparently need a little added incentive post-merger to attain an efficient equilibrium in their weakest-link choices.

CONCLUDING REMARKS

Section 3 has introduced you to the burgeoning field of behavioral game theory, a field that investigates how *Homo sapiens* play several of the classic games devised by game theorists to depict expected outcomes (i.e., analytical equilibria) when people interact in a variety of social situations. In particular, behavioral game theory identifies how *Homo sapiens* deviate from a game's analytical equilibrium and devises tweaks to the game in an effort to gain insight into what might be causing the deviation.

We began our exploration of the field in Chapter 7 by studying the classic bargaining games—Ultimatum Bargaining, Nash Demand, and Finite-Alternating Offers. The distinguishing features of these games are (1) they are played sequentially (i.e., one player makes a choice first and then the other player chooses); (2) the analytical equilibrium is premised upon each player thinking ahead about the other player's subsequent move before choosing what to do presently (i.e., each player thinks iteratively); and (3) the process of thinking iteratively requires that each player first considers what she should choose to do in the game's final stage, and then work backward to what to do in the game's initial stage. The resulting equilibrium is known as "subgame perfect" solved via "backward induction" and reflective of an "iterated dominance" in the player's respective strategies. As complicated as all of this sounds, we learned that *Homo sapiens* often achieve this type of equilibrium in a variety of contexts.

Likewise, we studied a variety of multiple-stage games with subgame perfect equilibria involving sequential moves by the players, each in their own way demonstrating the concept of iterated dominance. Recall the Escalation, Burning Bridges, Police Search, Dirty Faces, and Trust Games. In the Dirty Faces game, the trajectory toward the game's analytical equilibrium depends upon an initial, random draw of cards. One draw quickly leads players to the equilibrium; the other draw necessitates more intensive "iterated knowledge" on the part of the players. As expected, *Homo sapiens* do a better job of attaining the equilibrium when less iterated knowledge is required of them. In the Trust game, *Homo sapiens* players demonstrate excessive trust and insufficient trustworthiness.

Next, we considered a series of games in this chapter where players are again expected to use iterative thinking but this time in contexts where decisions are made simultaneously, not sequentially. As a result, players are expected to use "forward induction" in predicting how their opponents will play rather than backward induction. In some of these games, the analytical equilibria exist in "pure strategies," where the payoffs are structured in such a way as to make non-cooperative strategies "dominant" over others and the corresponding equilibria inefficient. This was the case with the famous Prisoner's Dilemma where, unfortunately, *Homo sapiens* tend to mimic self-interested *Homo economicus* and obtain the inefficient equilibrium. However, when *Homo sapiens* are allowed to play the game repeatedly, cooperation among players ensues.

Other simultaneous games do not exhibit dominant strategies, leading instead to "mixed-strategy" equilibria where players choose their strategies randomly but in accordance with some rule, such as flipping a fair coin. The Stag Hunt, Zero-Sum Game, and Battle of The Sexes are examples of games with mixed-strategy equilibria. Empirical research suggests that *Homo sapiens* are capable of attaining a games' mixed-strategy equilibrium, but only with the aid of additional incentives such as one- or two-way communication (known as "cheap talk") among players, or the availability of "outside options" for one of the players that the other player is also aware of. Similar results were found for other simultaneous-move games such as the Market Entry and Weakest Link Games. In the Market Entry Game, *Homo sapiens* responded to both previous-round and cumulative feedback by more closely mimicking the analytical equilibrium. In Weakest Link Games, smaller player groups achieved the analytical equilibrium more often. But when players played two games simultaneously—each game with a different set of players—the analytical equilibrium was obtained less often.

Therefore, it appears that *Homo sapiens* tend to perform more like *Homo economicus* in social settings (i.e., in games played with multiple players) than in individual settings (e.g., the laboratory experiments presented in Section 2), particularly when appropriate incentives are provided in the social setting. *Homo sapiens* tend to learn in both settings through repeated play and with appropriate incentives. For example, with appropriate incentives, if the opt-out payoffs are set too low or too high, players with access to opt-out options are more likely to choose strategies that deviate from the analytical equilibrium. This has been demonstrated in Stag Hunt games. As we learned from experiments with Ultimatum Bargaining and the Beauty Contest, the level of a game's stakes itself does not generally influence the probability that players will attain an analytical equilibrium.

As we proceed to Section 4, recall that the experiments and games we have studied in Sections 1–3 have traditionally been conducted with relatively small samples of university students. As such, the generalizability of their results to wider populations is justifiably drawn into question. In Section 4, we explore empirical studies based upon larger and more diverse (i.e., more "representative") samples of individuals, the results from which are, by design, more generalizable to wider populations.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013), those marked with a "‡" are adopted from Cartwright (2014), and those marked with a "≬" are adopted from Dixit and Nalebuff (1991).

1. † Consider the Prisoner's Dilemma game depicted below in which two prisoners are accused of a crime. Both are isolated in the prison. Without a confession, there is not enough evidence

to convict either. A prisoner who confesses will be looked upon with lenience. If one prisoner confesses and the other does not, the prisoner not confessing will be imprisoned for a much longer time than if she had confessed. The payoffs for the prisoners are depicted in the matrix below (Prisoner 1's payoffs are to the right of the comma in each cell, and Prisoner 2's are to the left of the comma in each cell). (a) If a selfish prisoner plays this prisoner's dilemma against an opponent she believes to be altruistic, what will her strategy be? (b) Now suppose the prisoner's dilemma is played three times in sequence by the same two prisoners. How might a belief that the other prisoner is altruistic affect the play of the selfish prisoner? Is this different from your answer to part (a)? What has changed?

		Prisoner 1	
		Confess	Don't Confess
Prisoner 2	Confess	- 10 , - 10	60 , - 100
	Don't Confess	- 100 , 60	50 , 50

2. In discussing strategies for a finitely repeated Prisoner's Dilemma game, it was mentioned that "tit-for-tat" may be a strategy that induces cooperation among players in an infinitely repeated version of the game. Recall that a player adopts a tit-for-tat strategy when he cooperates in the first round of the game and then, in each successive period, mimics the choice (deviate or cooperate) made by his opponent in the previous period. Thus, if Player 1 plays tit-for-tat and Player 2 deviates in the first round, then Player 1 commits to deviating in the second round, and so forth. If Player 1 believes that Player 2 is also playing tit-for-tat, and the game's payoff matrix is the same as that originally presented in this chapter (repeated below for reference), is it still in Player 1's interest to play tit-for-tat? Explain.

		Player 2	
		Cooperate	Deviate
Player 1	Cooperate	4 , 4	1 , 6
	Deviate	6 , 1	2 , 2

3. ‡ Consider the game below, played between Alan and Emma. Alan's payoffs are represented by

the first dollar value in each cell of the matrix, and Emma's payoffs are represented by the second dollar value. Identify each of this game's equilibria. Which equilibrium would you prefer Alan and Emma to reach together? What incentive might you give Emma and Alan to ensure that they would reach the preferred equilibrium?

		Emma	
		Low Effort	High Effort
Alan	Low Effort	\$7, \$7	\$7, \$0
	High Effort	\$0 , \$7	\$13, \$13

4. Choose a simultaneous game you learned about in this chapter. In what way would permitting communication between the players before the game begins affect the game's outcome?

5. In the discussion of the Public Good game, it was mentioned that a Provision-Point Mechanism has been found in experiments to increase contributions to public goods. Explain why this is the case.

6. ‡ Describe the similarities and differences between the Weakest Link and Threshold Public Good games, where the threshold in the public good game is a provision-point mechanism.

7. Is reaching an international agreement on the control of climate change similar to a Stag Hunt game? Explain.

8. Suppose Joe and Donald face the Zero-Sum game presented below. Calculate the game's mixed-strategy equilibrium.

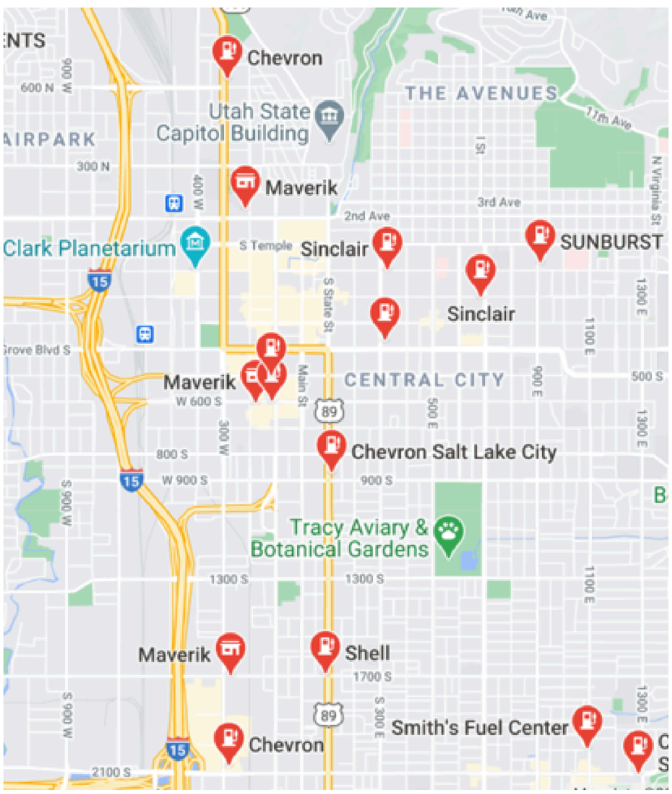
		Joe	
		Left	Right
Donald	Up	3, -3	-2 , 2
	Down	-1 , 1	0, 0

9. Recall that in Cooper et al.'s (1989) Battle of the Sexes experiments with *Homo sapiens*, when

one of the two players is allowed to communicate with the other player (i.e., there is "one-way communication") the players coordinate their choices 96% of the time! However, with simultaneous two-way communication between the two players, they coordinate only 42% of the time! Explain what happened.

10. We demonstrated how to solve for the Penalty Kick game's mixed-strategy equilibrium. Suppose you were new to the game of soccer (or football) and assigned to play the goalie position. After watching the following YouTube video, what strategy might make the most sense for you to adopt on penalty kicks: https://www.youtube.com/watch?v=3yWZZR9ZodI.

11. The map below identifies (with red markers) the locations of gas stations in Salt Lake City, Utah (Utah's capital city). Do these gas station locations depict a pure strategy equilibrium for the Hotelling Game? Explain.



Source: Google Maps

12. In this chapter, we learned that when an individual acquires private information about something, this added information does not necessarily make the individual better off. In particular, when an individual (say, Player 1) acquires private information about something of common interest to both himself and another individual (say, Player 2), and Player 2 knows Player 1 has acquired this private information, Player 1 could actually be made worse off as a result of Player 2 changing her strategy in response to the fact that she knows Player 1 now has additional information. Whew! Can you think of a real-life example where the acquisition

of private information actually makes the now-better-informed individual worse off? For inspiration in formulating your answer, consider watching this trailer of the 2019 movie *The Farewell*: https://www.youtube.com/watch?v=RofpAjqwMa8.

13. Analyze this excerpt from the British game show *Golden Balls* in the context of something you have learned about in this chapter: https://www.youtube.com/watch?v=S0qjK3TWZE8.

14. ◊ This question demonstrates how one player can turn a disadvantage encountered in a simultaneous-move game into an advantage by making a prior public announcement and thereby transforming the simultaneous-move game into a sequential-move game. To show this, suppose the players' payoff matrix is as shown below. (a) Determine this game's analytical equilibrium in pure strategies (i.e., its PSE). (b) Explain why Player 1 is not happy with this equilibrium. (c) Suppose Player 1 decides to preempt this game's equilibrium outcome by announcing what his effort level will be before Player 2 decides on her effort level. Describe this new game in its extensive form (i.e., in the form of a decision tree). Using iterated dominance, identify the equilibrium for this game. Did Player 1's preemptive announcement work to his advantage (i.e., did Player 1 give himself a first-mover advantage)? Explain. (d) Explain why Player 1's preemptive announcement does not necessarily have to be considered credible by Player 2 to have its beneficial effect on Player 1.

		Player 2	
		Low Effort	High Effort
Player 1	Low Effort	4 , 3	2 , 4
	High Effort	3 , 2	1 , 1

15. ◊ In 1944, the Allies were planning their liberation of Europe while the Nazis planned their defense. The Allies were considering two possible landing sites—the beaches of Normandy or Pas de Calais. Calais was considered more difficult to invade but more valuable to win given its proximity to Belgium and Germany. Suppose the payoff matrix facing the Allies and the Nazis is as depicted below. (a) What type of game does the payoff matrix represent? (b) Calculate the game's mixed-strategy equilibrium. Is this equilibrium consistent with what actually occurred in 1944? Explain. (c) Calculate the Allies' expected payoff from following its equilibrium strategy.

		Nazi Defense	
		Normandy	Calais
Allied Landing	Normandy	60 , - 60	80 , - 80
	Calais	100 , - 100	20 , - 20

16. Suppose two players are involved in a game where one of the player's payoffs (both when she cooperates and deviates) are much larger than the other player's (e.g., because the player with the larger payoffs (say, Player 1) is a much larger producer than Player 2). The payoff matrix for this game is provided below. Determine this game's analytical equilibrium. Comment on this equilibrium in relation to the standard analytical equilibrium obtained in the Prisoner's Dilemma.

		Player 2	
		Cooperate	Deviate
Player 1	Cooperate	64 , 16	48 , 24
	Deviate	60 , 12	40 , 16

Media Attributions

- Figure 1 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 2 (Chapter 8) © Fehr and Gachter; 2000 American Economic Association is licensed under a All Rights Reserved license
- Figure 3 (Chapter 8) © Fehr and Gachter; 2000 American Economic Association is licensed under a All Rights Reserved license
- Figure 4 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 5 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 6 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 7 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 8 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license

- Figure 9 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 10 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Hotellings Game Figure © Arthur Caplan is licensed under a CC BY (Attribution) license
- Is More Information Always Better Figure © Arthur Caplan is licensed under a CC BY (Attribution) license
- Is More Information Always Better Figure (second figure) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Weakest Link Figure © Arthur Caplan is licensed under a CC BY (Attribution) license
- Weakest Link With Local Interaction Figure v1 © Arthur Caplan is licensed under a CC BY (Attribution) license
- Weakest Link With Local Interaction payoff matrix v2a © Arthur Caplan is licensed under a CC BY (Attribution) license
- Weakest Link With Local Interaction eight player schematic © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 1 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 3 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 8 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 12 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 14 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 15 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure for Study Question 16 (Chapter 8) © Arthur Caplan is licensed under a CC BY (Attribution) license

PART IV.

SECTION 4 - EMPIRICAL STUDIES AND FIELD EXPERIMENTS

In Sections 1–3, we learned something about how the field of behavioral economics has evolved: from laboratory experiments undertaken with individuals and games played by groups of individuals to revised theories that canonize the specific ways in which *Homo sapiens'* choice behavior deviates from the rational-choice model of *Homo economicus*. In Section 4 we consider novel empirical studies and field experiments that have been conducted in a myriad of settings, some of which test (with larger samples of *Homo sapiens*) the main tenets of behavioral economic theory. In particular, they test our penchants for loss aversion, reference dependence, time inconsistency, and generally our reactions to how choice situations are framed (e.g., narrowly versus broadly). These studies seek answers to the same types of questions the laboratory experiments and games of the previous two sections posed to economists and psychologists; how exactly do *Homo sapiens*' choice behaviors deviate from those of *Homo economicus* in a variety of real-life situations that often have pronounced social consequences? What might explain these deviations?

Other empirical studies conducted by behavioral economists test for evidence of sub-optimal behavior among *Homo sapiens* that expresses itself as discrimination, racial bias, corruption, crime, waste, homelessness, drunk driving…I think you get the point. These behaviors manifest themselves as social ills. Less severe forms of sub-optimal behavior (e.g., underperforming students and teachers, insufficient savings and investment by households and businesses, tax evasion, and low voter turnout) have also received attention from behavioral economists. These types of behaviors lead scholars and policymakers to ask what social constructs (or, what Thaler and Sunstein (2009) call choice architecture) might be developed to mitigate the negative social consequences associated with these types of sub-optimal behavior.

Referring to the diagram presented in the This Book's Approach section, Section 4 pertains to the diagram's middle portion, where we investigate empirical studies that have been undertaken in the real world and the corresponding choice architecture that has been developed—studies and architecture spawned by the now famous experiments, games, and concomitant theory of behavioral economics.

For the most part, these empirical studies are based not on experiments run in laboratories but rather on experiments run in the "fields" of our daily lives. We call them "natural experiments" or "field experiments." The findings we obtain from these studies can inform public policy and point us (or, as Thaler and Sunstein would say, nudge us) toward crafting more enlightened policies that reflect and channel our idiosyncrasies, biases, foibles, and social norms (i.e., our humanity) for improved social outcomes. As you proceed through what is essentially a menagerie of different behavioral economics case studies, bear in mind that one of the goals of this section is to introduce you to research methods as well as attendant empirical results. Along the way, you will be introduced to the myriad approaches researchers use to present their statistical results.

Note that, unlike a normal textbook chapter, this section is not necessarily meant to be read from beginning to end. There are too many disparate studies included herein, with no particular theme to tie them all together. Rather, only those studies deemed relevant for your course—whether this relevance is predetermined by the course's instructor or the student's personal judgment—are meant to be chosen from the wide universe of studies compiled below, and conceivably linked with the concepts and topics from Sections 1 – 3 that were previously covered in the course. The Linkages Matrix contained in Appendix E can help you with making these linkages.

Media Attributions

CHAPTER 9.

THE STUDIES AND EXPERIMENTS

DISCRIMINATION IN THE WEAKEST LINK GAME SHOW

We begin with a study of potential discrimination exhibited by contestants in a popular British game show called the *Weakest Link*. This is not a show where contestants play the Weakest Link Game described in Chapter 8. Rather, the goal of the game is for a group of contestants to vote individuals off the show one-by-one in successive rounds until only two contestants remain to compete for a grand prize. The contestants who are voted off are considered "weak links" as a consequence of strategic play by the individual members of the group. Weak links are considered liabilities to the remaining group members in terms of building up the jackpot and, potentially, being one of the two remaining contestants to play for it. If you have the time and are interested in watching the show, check out this Youtube video (and hold onto your Bowler hat. It's a lively, fast-paced contest).

Levitt (2004) observed that contestant voting behavior on the show provides an opportunity to distinguish between what he calls taste-based (bad!) and information-based (not as bad!) theories of discrimination. Taste-based discrimination occurs when an individual prefers not to interact with a particular class of people, and he is willing to pay a financial price to avoid such interactions. In contrast, an individual practicing information-based discrimination has no animus against a particular class of people but discriminates nonetheless because she has less reliable (i.e., noisy) information about them.

Contestants answer trivia questions over a series of rounds, and one contestant is eliminated each round based upon the votes of the other contestants until only two contestants remain. The last two contestants compete head-to-head for the winner-take-all prize. Because the prize money at stake is potentially large (the money is an increasing function of the number of questions answered correctly by the group over the course of the game's rounds), contestants have powerful incentives to vote in a manner that maximizes their individual chance of being one of two remaining contestants to compete for the jackpot.

In the early rounds of the contest, strategic incentives encourage voting for the weakest competitors. However, in later rounds, the incentives reverse, and the strongest competitors become the logical target of eviction. Both theories of discrimination suggest that, in early rounds, excess votes will be cast against people targeted for discrimination. If group members practice taste-based discrimination, then in later rounds, these excess votes would persist, whereas if information-based discrimination is practiced, then votes against the targeted people would diminish.

Levitt (2004) found that contestants voted strategically in early rounds of the game but not in later rounds. Specifically, voting strategically means both voting off players who, in the game's early rounds, more frequently answer questions incorrectly or "take a pass" on providing an answer (and thus do not contribute as much toward the ultimate jackpot by answering correctly), *and* voting off players in later rounds who consistently answered questions correctly in the previous rounds (since they now present more of a threat to make it to the game's final round). There is little evidence to suggest that contestants discriminate against women, Hispanics, and people of African descent. However, some evidence suggests taste-based discrimination against older players.

For those of you with a background in statistics or econometrics, Levitt's (2004) specific results are presented in the table below:

REGRESSION ANALYSIS OF VOTES RECEIVED

	EARLY ROUNDS (N = 1,191)		MIDDLE ROUNDS (N = 1,599)		FINAL ROUND (N = 483)	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Female	−.09 (.09)	−.07 (.08)	.00 (.06)	−.04 (.05)	−.06 (.08)	−.10 (.08)
Black	.13 (.11)	.07 (.09)	−.05 (.08)	−.00 (.06)	−.01 (.09)	−.01 (.09)
Asian	−.34 (.17)	−.20 (.13)	.28 (.19)	.33 (.15)	−.17 (.24)	−.13 (.26)
Hispanic	.32 (.29)	.17 (.24)	.40 (.25)	.41 (.19)	−.34 (.28)	−.35 (.24)
Age 50+	.34 (.19)	.17 (.17)	.30 (.13)	.30 (.11)	.49 (.16)	.42 (.16)
% Correct this round (deviation from other players)	. . .	−2.44 (.11)	. . .	−1.71 (.09)	. . .	−.47 (.16)
Squared % correct this round (deviation from other players)	. . .	2.53 (.30)	. . .	.73 (.25)	. . .	.24 (.37)
Cumulative % correct (deviation from other players)	. . .	−.90 (.38)	. . .	−.70 (.14)	. . .	−.33 (.33)
Squared cumulative % correct (deviation from other players)	. . .	1.37 (.88)	. . .	.59 (.33)	. . .	2.77 (1.47)
Cumulative opponents voted against who are still alive	. . .	.12 (.29)	. . .	.27 (.04)	. . .	.09 (.04)
R^2	.013	.387	.010	.257	.043	.089
Education and region dummies included?	No	Yes	No	Yes	No	Yes

NOTE.—The dependent variable in all columns is the number of votes received by a contestant in a given round. The unit of observation is a contestant round. Estimation is done with ordinary least squares. Standard errors, clustered by episode and round, are in parentheses. All regressions include an exhaustive set of interactions controlling for round by show-length interactions. The even columns include state-of-residence fixed effects and occupation dummies, but these coefficients are not reported in the table. Early rounds correspond to the first two rounds of the prime-time show (1 hour long) and the first round of the daytime show (30 minutes long). Middle rounds correspond to rounds 3–5 of the -time show and rounds 2 and 3 of the daytime show. Final round corresponds to round 6 of the prime-time show and round 4 of the daytime show.

(Levitt 2004)

In a proverbial nutshell, the numbers not in parentheses indicate the sign and size of a given variable's effect on receiving votes (in favor of being removed from the group). The numbers in parentheses are called "standard errors" (SE). They are strictly positive numbers. Roughly speaking, the smaller an SE relative to the magnitude of its corresponding variable's effect on votes received, the more "statistically significant" is the effect. For example, consider the effect of being female in the game's early rounds. Although this effect is negative (-0.09), because the magnitude is very close (in this case exactly equal) to its corresponding SE of 0.09, we say that the 'female effect' is non-existent in a statistical sense.[1]

Taste-based discrimination is evident when an effect is positive and statistically significant in both the early and later rounds. Especially in the early rounds, votes should be based strictly on performance (i.e., a contestant's ability to answer questions correctly on behalf of the group), not gender, ethnicity, or age. We see from column (1) that solely older contestants (age 50+) satisfy this condition (in column (1) the Age 50+ value of 0.34 is almost double the size of its corresponding SE of 0.19).[2] Results in columns (3) and (4) for Age 50+ suggest a positive Age 50+ effect in the middle rounds as well.

Lastly, results for the variable "% Correct this round" provide evidence of the previous claims that contestants vote strategically in the early rounds of the contest but not in the later rounds. The large negative (and statistically significant) effect in column (2) of -2.44 indicates that, all else equal, contestants cast fewer votes for fellow contestants who provide correct answers more often. In the middle and final rounds, this effect should become positive if contestants vote strategically. We see

1. Roughly speaking, an effect is statistically significant when its corresponding SE is half of the size the magnitude of the effect or less.
2. The fact that the effect of being Asian is negative and statistically significant suggests "reverse discrimination" is potentially being practiced by contestants.

from columns (4) and (6) that this does not happen—large negative (and statistically significant) effects persist in these later rounds.

DISCRIMINATION IN PEER-TO-PEER LENDING

Pope and Sydnor (2011) also test for discrimination in a novel context—peer-to-peer lending on the website Prosper.com. Peer-to-peer lending is an alternative credit market that aggregates small amounts of money provided by individual lenders to fund moderately sized, uncollateralized loans to individual borrowers. Like most standard credit applications, Prosper.com publicizes loan information from the prospective borrower's credit profile. However, borrowers may also include optional personal information in their listing in the form of photographs and text descriptions. These pictures and descriptions can provide potential lenders with signals about characteristics such as race, age, and gender that anti-discrimination laws typically prevent traditional lending institutions from using.

Using data from 110,000 loan listings appearing on Prosper.com from 2006-2007, the authors find evidence of significant racial discrimination in this market. Loan listings that include a photograph of a Black borrower result in a 30% reduction in the likelihood of that loan receiving funding, all else equal. Further, a loan listing tied to a Black borrower results in an interest rate that is 60 basis points higher than an equivalent listing for a white borrower.[3] These results meet what the authors claim is a necessary condition for taste-based discrimination as defined by Levitt (2004). The question of whether the sufficient condition for this type of discrimination is met depends upon whether Black borrowers have statistically lower loan default rates and produce higher net returns for lenders. If the answer is "yes," then together with the fact that they are less likely to receive funding and pay higher interest rates on loans they do receive, evidence of taste-based discrimination against Black borrowers is evinced.[4]

The authors find that Black borrowers are approximately 36% more likely to default on their loans than are Whites with similar characteristics, and a lender's average net return from a loan to a Black borrower is eight percentage points lower over a three-year period. Thus, they conclude that discrimination in peer-to-peer lending could in fact be information- rather than taste-based.

WHAT'S IN A NAME?

According to the Economic Policy Institute's (EPI's) recent assessment of the US labor market, Black workers are twice as likely to be unemployed as White workers overall (6.4% vs. 3.1% unemployment rates, respectively)—a gap that, while narrower, persists among workers Black versus White workers with college degrees (3.5% vs. 2.2%) (Williams and Wilson, 2019). Further, when employed, Black workers with college or advanced degrees are more likely than their White counterparts to be underemployed—roughly 40% of Black college graduates are in jobs that typically do not require a college degree, compared with only 31% of their White counterparts. The EPI concludes that

3. Pope and Sydnor also find less-statistically significant discrimination against older and overweight borrowers, as well as borrowers who either did not provide a photograph of themselves or look unhappy in the photo provided. To the contrary, the authors find discrimination in favor of women and military veterans.

4. Chiang and Wainwright (2004) provide nice examples of necessary and sufficient conditions in a variety of contexts. My favorite is the fact that being male is a necessary condition for being a father (not vice-versa). A male who fathers a child has met both the necessary and sufficient conditions for being a father (at least biologically speaking).

persistence in relatively high Black unemployment and skills-based underemployment indicates that racial discrimination remains a failure of the US labor market, even when the market is tight.

The question naturally arises as to whether employers do in fact favor White applicants over similarly skilled Black applicants (i.e. do employers discriminate among job candidates based upon race?). Bertrand and Mullainathan (2004) provide an answer based upon an intriguing field experiment where fictitious resumes were sent by the authors in response to help-wanted ads in Boston and Chicago newspapers. To manipulate perceived race, the resumes were randomly assigned Black- or White-sounding names, such as Lakisha Washington or Jamal Jones (Black-sounding names) in response to half the ads, and Emily Walsh or Greg Baker (White-sounding names) in response to the other half.

Because they were also interested in how credentials affect the racial gap in interview callbacks, Bertrand and Mullainathan (2004) varied the quality of the resumes. Higher-quality applicants had on average more labor-market experience and fewer holes in their employment history. These applicants were also more likely to have an email address, have completed some certification degree, possess foreign language skills, or have been awarded some honors. The authors generally sent four resumes in response to each ad—two higher-quality and two lower-quality. They randomly assigned Black-sounding names to one of the higher- and one of the lower-quality resumes. In total, the authors responded to over 1,300 employment ads in the sales, administrative support, clerical, and customer services job categories, sending out nearly 5,000 resumes in total.

Overall, White names received 50% more callbacks for interviews, which Bertrand and Mullainathan (2004) translate into a White name being as valuable as an additional eight years of experience on a Black person's resume. Callbacks were also more responsive to resume quality for White than for Black names. The racial gap in interview callbacks was uniform across occupation, industry, and employer size. The authors also found that living in a "better" neighborhood (wealthier or more-educated or Whiter) increased callback rates. However, Blacks were not helped more than Whites by living in better neighborhoods. As the authors point out, if ghettos and bad neighborhoods are particularly stigmatizing for Blacks, one might have expected Blacks to have been helped more by having a better address. These results do not support this hypothesis.

Bertrand and Mullainathan (2004) also find that, across all sent resumes, the difference in percent callbacks for White- versus Black-sounding names is a statistically significant 3.2%. Callback discrimination based upon race occurs against both men and women, and the discrimination against Black women occurs mostly in conjunction with administrative rather than sales jobs.[5]

It is humbling to think that *Homo economicus* employers, whose color-blindness is a patent feature of their rational minds, would not fall victim to racial discrimination in the hiring process. Thankfully, as the topic below, *Awareness Reduces Racial Discrimination*, suggests, when racial discrimination is brought to light its practice tends to dissipate.

CAN LOOKS DECEIVE?

Similar to Prosper.com, but with a bit more (how shall we say?) gravity, online dating services create a natural setting within which to assess the impacts of a person's physical appearance on a transaction between that person and a potential (how shall we say?) customer. One dating site, OkCupid, recently

5. Bellemare et al. (2023) investigate similar patterns of discrimination among people with physical disabilities in a large-scale field experiment. They find that roughly 50% of private firms discriminate against people with physical disabilities. However, on average callback rates for disabled individuals is double that of the non-disabled.

became interested in answering the simple question, to what extent do looks deceive? The site's answer is based upon a natural experiment conducted with their users on what OkCupid named "Love is Blind Day," January 15, 2013, celebrating the release of their new phone app. Comparing that day's messaging data to the average day's historically, Rudder (2014) uncovered several interesting (how shall we say?) relationships in the data.

For example, OkCupid's site metrics (number of new conversations started per hour) were far beneath a typical Tuesday's during the peak hours of 9 a.m. to 4 p.m. It seems that without the ability to view a prospective date's photo, users were less motivated to make an initial inquiry. Nevertheless, Rudder reports that the conversations initiated during these seven hours without photos went deeper and contact details (e.g., email addresses and phone numbers) were exchanged more quickly. Sadly though, when the photos were restored at 4 p.m. sharp, the 2,200 users who were in the middle of their conversations that had started "blind" dissipated. As Rudder puts it, restoration of the photos was like turning on the bright lights at the bar at midnight. Conversations that had consisted of two messages prior to the 4 p.m. bewitching hour witnessed the largest drop relative to normal.

Curious about the extent to which a person's photo matters on OkCupid, Rudder performed a simple test based upon a randomly chosen subsample of users. Half of the time their pages were accessed by a prospective suitor, their profiles were kept hidden. And half of the time the profiles were not hidden. This generated two independent sets of ratings for each member of the sample—one rating when the picture and profile text were presented together, the other for when the picture was presented alone. Rudder found a strong positive correlation between the ratings with and without the profile text included, suggesting that a picture really is worth a thousand words. A person's rating was driven by the appeal of their picture rather than their profile. To put it less sanguinely, we *Homo sapiens* tend to be superficial when it comes to choosing our dating partners.[6]

THE SPILLOVER OF RACIALIZATION

To what extent might racial prejudice spill over into (i.e., infect) opinions about public policy, such as health care and fiscal stimulus? The election of Barack Obama in 2008 as the 44th President of the United States helps provide an answer. Using data from a nationally representative survey experiment, Tesler (2012) documents the impact of race and racial attitudes on opinions concerning national healthcare policy before and after Obama's election. The authors find that racial attitudes were both an important determinant of White Americans' opinions about healthcare policy in the fall of 2009 and that the influence of these attitudes increased significantly after President Obama became the face of the policy. Results from the experiment show that racial attitudes had a significantly greater impact on healthcare opinions when framed as part of President Obama's plan than they had when the same policies were attributed to President Clinton's 1993 healthcare initiative. In other words, Tesler uncovers what he calls a spillover of racialization, which situates Obama's race—and

6. These Appearance Effects uncovered by Rudder (2014) on OkCupid hound academic economists in their profession as well. Hale et al. (2021) find robust evidence that physical appearance has predictive power for the job outcomes and research productivity of PhD graduates from ten of the top economics departments in the US. Attractive individuals are more likely to study at higher-ranked PhD institutions and are more likely to be placed at higher-ranking academic institutions not only for their first job, but also for jobs as many as 15 years after their graduation. More attractive economics PhD graduates also have their published research cited more often by other researchers. On the flip side of this phenomenon, New York Times opinion columnist David Brooks says it best: "We live in a society that abhors discrimination on the basis of many traits. And yet one of the major forms of discrimination is lookism, prejudice against the unattractive. And this gets almost no attention and sparks little outrage" (Brooks, 2021; italics added).

the public's race-based reactions to him—as the primary reason why public opinion about national healthcare policy racialized in the fall of 2009.

As Tesler points out, spillover of racialization—whereby racial attitudes have a bearing on political preferences—is rather straightforward for race-targeted public policies such as affirmative action and federal aid to minorities. These types of issues are thought to readily evoke racial predispositions since a natural associative link exists between policy substance and feelings toward the groups who benefit from them. However, this link is not as readily apparent for broader issues such as healthcare and fiscal stimulus.

Tesler further avers that, after receiving little media attention during the first half of 2009, the debate over healthcare reform became one of the most reported news stories in America from early July through the remainder of the calendar year, so much so that roughly half of Americans reported following the healthcare reform debate very closely in 2009 (Pew Research Center, 2009). If, as the spillover of racialization hypothesis contends, Obama's connection to the issue helped racialize their policy preferences, then the effect of racial attitudes on White Americans' opinions should have increased from before to after his healthcare reform plan was subjected to such intense media scrutiny.

Tesler utilizes observational data from repeated cross-sectional surveys conducted by the American National Election Study (ANES). The ANES healthcare question asks respondents to place themselves on a seven-point government-to-private insurance preference scale. To obtain corresponding information on racial resentment, Tesler re-interviewed individuals who had participated in the ANES survey both before and after President Obama's election. The author argues that his racial-resentment measure taps into subtle hostility among White Americans toward Black Americans. The measure is based upon four questions about Black work ethic, the impact of discrimination on Black American advancement, and notions of Black people getting more than they deserve—themes thought to undergird a symbolic racialism belief system—and coalesced into a seven-point scale from low to high levels of racial resentment.

Tesler finds that, for White respondents, moving from those harboring the least amount of racial resentment to those harboring the most resentment increased the proportion of those saying that the national healthcare system should be left up to individuals by approximately 30 percentage points (from 10% to 40%) in December 2007, when President Clinton was the face of national healthcare policy. However, the same change in these individuals' resentment levels (i.e., again moving from those White respondents harboring the least amount of racial resentment to those harboring the most resentment) increased their support for private insurance by roughly 60 percentage points (from 10% to 70%) in November 2009, when President Obama served as the face of the same national healthcare policy—a statistically significant difference. This leads Tesler to conclude that with the election of President Obama racial attitudes became more important in White Americans' beliefs about healthcare relative to nonracial considerations like partisanship and ideology.

In an additional experiment, Tesler investigated opinions regarding the $787 billion economic stimulus package passed by Congress in 2009. Respondents were divided into two subsets. In one subset, respondents were asked if they thought the stimulus package approved by congressional Democrats was a good or bad idea, the other was asked the same question but with approval instead being granted by President Obama. The author finds that moving from least to most racial resentment decreased the proportion of White respondents saying that the stimulus program was a good idea by less than 10 percentage points when congressional Democrats are identified as the approving authority, but by approximately 70 percentage points when President Obama is identified as the approving authority. In other words, the incidence of racialization spillover is even more profound than it was regarding national healthcare policy.

AWARENESS REDUCES RACIAL DISCRIMINATION

In situations where racial discrimination is known to exist, does informing the public of its existence encourage perpetrators to repudiate its practice? Pope et al. (2018) devised a novel approach to answer this question. In 2003, the authors began by analyzing data from the National Basketball Association (NBA) for the years 1991-2003. They found that White and Black players received relatively fewer personal fouls when more of the referees officiating the game were of their own race. This in-group favoritism (or, alternatively stated, out-group racial bias) displayed by NBA referees was large enough to influence game outcomes.

In May of 2007, the results of this study received widespread media attention—front-page coverage in the *New York Times* and many other newspapers, and extensive coverage on major news networks, ESPN, and talk radio. Subsequently, the authors analyzed NBA data for the years 2007–2010 and found an absence of this out-group racial bias, although other biases were found to persist (e.g., referees tend to favor the home team, the team that is losing in a given game, and teams that are losing the game count in a playoff setting).

The table below contains Pope et al.'s (2018) specific findings:

	Pretreatment		Posttreatment	Change in bias
Sample	Original study 1991–2003	Out-of-sample 2004–2006	Out-of-sample 2007–2010	From 2004–2006 to 2007–2010
Out-group bias (extra fouls per 48 minutes when refereed by an out-group crew, relative to an in-group crew)	0.192^{***} (0.059)	0.214^{**} (0.107)	−0.0002 (0.089)	0.214^{**} (0.105)
Change in out-group bias, subsequent to "treatment"				−0.214 (0.139)
N	282,175	70,465	94,682	165,147
Sample mean	4.44	4.46	4.17	4.30

Notes. The years in each column refer to the year in which the season started. Each regression includes player and referee fixed effects and controls for home team and whether the player is a starter. Each observation is weighted by the number of minutes the player was in the game. Standard errors are in parentheses.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

(Pope, et al. 2018)

Similar to the presentation of Levitt's (2004) results (see *Discrimination in The Weakest Link Game Show* above), marginal effects are presented with their corresponding standard errors in parentheses. Pope et al. provide additional notation to distinguish statistically significant effects from those that are not—superscripts with more asterisks indicate more statistical significance; those effects without any asterisks indicate no statistical significance. The marginal effects in the pretreatment period—based upon data from the original study and an additional study covering the years 2004-2006—are positive and statistically significant (0.192 and 0.214, respectively), leading the authors to conclude that, prior to media attention, significantly more fouls were called on Black players when the referee crew was predominantly White. In the post-treatment period—based upon data from 2007-2010—the marginal effect is not statistically significant (i.e., no effect exists). Hence, the prior racial bias among referee crews in the NBA dissipated after having received widespread media attention. That's a slam dunk for the NBA and another one for *Homo sapiens* in general! Raising awareness of racial discrimination, especially when we can quantify its presence, serves as a nudge toward racial equality.

IMPROVING STUDENT PERFORMANCE

Levitt et al. (2016) designed a field experiment to test the effects of different incentive mechanisms on the academic performance of students in low-performing elementary, middle, and high schools in the Chicago public school system and, in the process, test for the existence of loss aversion and time inconsistency among the students. Students were offered one of the following rewards for improving upon a previous (baseline) computerized reading or math test: $10 in cash ("financial low"), $20 in cash ("financial high"), or a trophy and posting of a student's photograph in the school's entrance ("nonfinancial").

To test for loss aversion among the students, financial and non-financial rewards were delivered in one of two ways: (1) the test administrator held up the $10 bill, $20 bill, or trophy at the front of the room before the test began (the authors call this the "gain condition"), or (2) students received the $10 bill, $20 bill, or trophy at the start of the testing session and were informed that they would keep the reward if their performance improved and lose the reward if it did not ("loss condition"). The following results were obtained:[7],[8]

1. The $20 incentive (framed either as a gain or loss) delivered immediately after students completed the test increased the average student's test score. The $10 incentive did not increase the average student's test score and even lowered performance on future tests.

2. The trophy delivered immediately after the test increased the average student's test score less dramatically than the $20 incentive. Scores increased most dramatically for younger students who received trophies.

7. In an innovative field experiment conducted in Columbia, Barrera-Osorio et al. (2008) studied the effects of conditional cash rewards on student attendance and graduation rates (as well as contingent intra-family and peer-network dynamics). The authors' experiment consisted of three treatments: a basic conditional cash transfer treatment based upon school attendance, a savings treatment that postponed the bulk of the conditional cash transfer to just before the student was scheduled to re-enroll in school, and a tertiary treatment where a portion of the cash transfer was made conditional upon a student's graduation and tertiary enrollment. Barrera-Osorio et al. found that, on average, the combined cash incentives increased school attendance, pass rates, enrollment, graduation rates, and matriculation to tertiary institutions. Changing the timing of the payments (e.g., in the savings and tertiary treatments) did not affect attendance rates relative to the basic cash transfer treatment, but did significantly increase enrollment rates at both the secondary and tertiary levels. The tertiary treatment was particularly effective, increasing attendance and enrollment at secondary and tertiary levels more than the basic treatment. The authors also found some evidence that the cash transfers caused a reallocation of responsibilities within a student's household. Siblings (particularly sisters) of participating students worked more and attend school less than siblings of students who did not participate in the experiment. In addition, peer influences were relatively strong in influencing a student's attendance decisions.
8. Unrelated to loss-aversion treatments, but nevertheless of interest when it comes to the question of student performance, Gershenson et al. (2022) use data from a field experiment with K-3 public school students in Tennessee to test whether the teacher's race has an impact on performance. They find that Black students randomly assigned to at least one Black teacher are roughly 13 percent more likely to graduate from high school and 19 percent more likely to enroll in college compared to their Black schoolmates who did not study under at least one Black teacher. Black teachers have no statistically significant effect on White students' high school graduation rates or likelihood of enrolling in college.

3. The average student's test score increased more in the loss condition than in the gain condition, but the difference is not statistically significant. Hence, the average student does not exhibit loss aversion with respect to how the reward for improved performance is distributed.

4. Delayed rewards (delivered one month after completion of the exam rather than immediately) did not increase the average student's test score. This suggests the existence of hyperbolic discounting, where rewards delayed in the near term are discounted at an excessively high rate (recall our earlier exploration of this phenomenon in Chapter 4).

5. Overall, math scores increased more than reading scores across all students. Boys increased their scores in these subjects more than girls.

IMPROVING TEACHER PERFORMANCE

Fryer et al. (2022) demonstrate that, unlike Levitt et al.'s (2016) findings for elementary and middle school students in Chicago, exploiting the power of loss aversion—where the student's teachers are paid at the beginning of the school year and asked to give back the money if their students do not improve sufficiently (loss treatment)—leads to statistically significant increases in their students' math test scores. A second treatment identical to the loss treatment but with year-end bonuses linked to student performance (gain treatment) yields smaller and statistically insignificant results. The authors conclude that because teachers exhibit loss aversion (in terms of rewards tied to their students' academic performance), a loss-treatment approach is the most effective way to incentivize teachers to improve student performance.

In specific, Fryer et al. find that, all else equal, the average student who was taught by a teacher who had been randomly assigned to the loss treatment gained (statistically significant) percentile-ranking points relative to her nine nearest students during the math exam; gains that persisted in time after the treatment. Students who were taught by teachers who were randomly assigned to the gain treatment showed markedly lower and statistically insignificant gains. Therefore, it seems as though a teacher's performance can be more effectively nudged by appealing to his sense of loss aversion rather than merely the teacher's desire for gain.

HEALTHCARE REPORT CARDS

Recall from Chapter 8 the simultaneous-move game where provisioning one of two players with additional information actually perversely affected the game's analytical equilibrium. The message was clear. The rational choice model's tenet that more information leads to improved performance is not universal. Especially when it comes to the experience of *Homo sapiens*, situations where the provision of additional information leads to a perverse outcome are not necessarily in short supply.

Dranove et al. (2003) provide a seminal example with their study of Healthcare Report Cards—public disclosure of patient health outcomes at the level of the individual physician or hospital or both—that are intended to improve the performance of healthcare providers. In their study, the authors analyzed New York's and Pennsylvania's publications of physician and hospital coronary artery bypass graft (CABG) surgery mortality rates in the 1990s. At the time, the merits of these types of report cards were in much debate. Supporters argued that report cards enable patients to identify the best physicians and hospitals while simultaneously giving healthcare providers powerful incentives to improve quality. Skeptics countered that report cards encourage providers to "game" the system by avoiding sick patients and/or seeking healthy patients.

As Dranove et al. point out, low-quality providers have strong incentives to avoid the sick and seek the healthy under this type of reporting system. By shifting their practice toward healthier patients, inferior providers make it difficult for report cards to distinguish them from their higher-quality counterparts because relatively healthy patients have higher likelihoods of better outcomes regardless of provider. As the authors put it, low-quality providers can therefore pool with their high-quality counterparts, making it more difficult for the report cards to distinguish between the two.

Spoiler alert: The authors find that while the report card system increased the quantity of CABG surgeries among patients suffering from acute myocardial infarction (AMI) (i.e., heart attacks), it changed the surgery's incidence from sicker AMI patients toward healthier AMI patients. Overall, this led to higher costs and deterioration of outcomes, especially among the sicker AMI patients (i.e., the report cards were welfare-reducing).

Dranove et al. find that the introduction of report cards increased the probability that the average AMI patient would undergo CABG surgery within one year of hospital admission by between 0.60 to 0.91 percentage points. As the authors point out, these report-card effects are considerable, given that the probability of CABG within one year for an elderly AMI patient during their sample period was approximately 13%. However, the report-card effects did not occur immediately (i.e., within one day of admission to the hospital). Indeed, the immediate report-card effect is estimated to have been negative for the average AMI patient (ranging from -0.59 to -0.78 percentage points). The authors also find evidence to suggest that the report card system led to sicker patients being less likely to undergo CABG surgery within one year of admission. Report-card effects on average (1) led to increases in total hospital expenditures in the year after admission of an AMI patient, (2) provided some evidence of increased patient readmission with heart failure within one year, and (3) provided some evidence of increases in mortality within one year of admission. These perverse welfare effects were particularly strong among sicker AMI patients.

This is one of several examples in the empirical literature of perverse outcomes associated with what, on the surface, would seem to be naturally beneficial incentives, or nudges, meant to improve the social welfare of *Homo sapiens*, in this case with respect to health care.

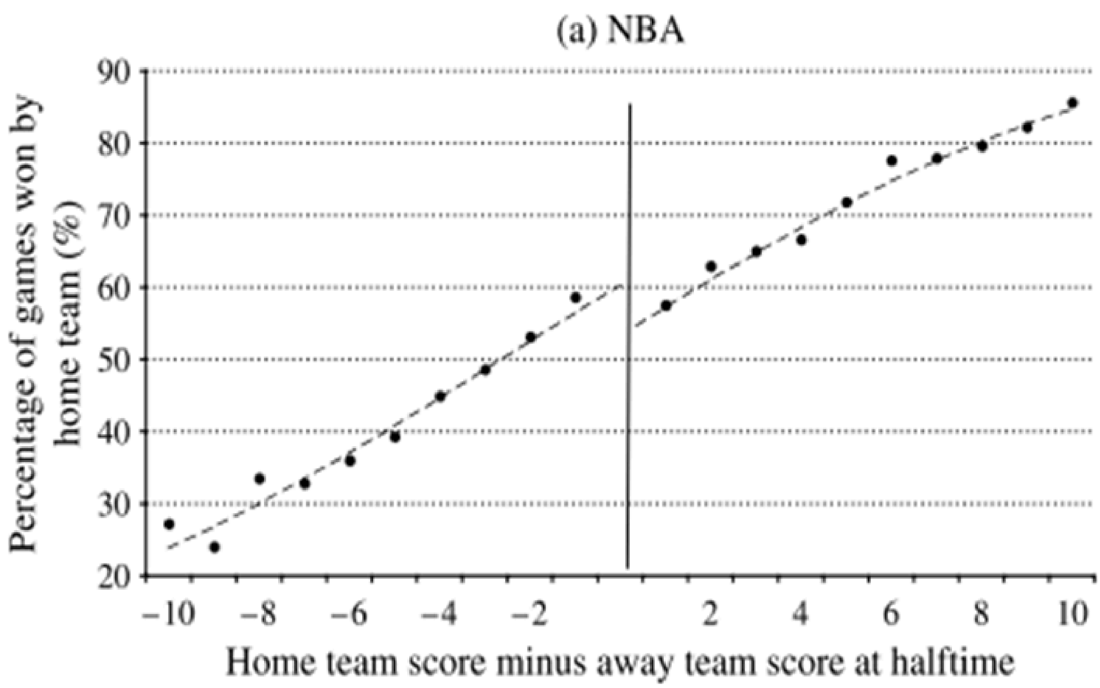
LOSING CAN LEAD TO WINNING

Berger and Pope (2011) conducted another study using data from the NBA, this time seeking to determine if teams who, going into halftime of a typical game and down a certain number of points, collectively exhibit loss aversion—in terms of not wanting to lose the game—by winning the game in the end. In other words, do NBA players demonstrate loss aversion collectively as a team?

The authors analyzed more than 18,000 NBA games played from 1993-2009 and found that teams behind by one point at halftime win more often than teams ahead at halftime by one point—approximately 6% more often than expected. This finding suggests the presence of (1) loss aversion—being behind at halftime motivates a team not to lose more than being ahead at halftime motivates a team to win, (2) diminishing sensitivity—the losing team cannot be too far behind at halftime, and (3) reference dependence—being behind at halftime helps the losing team establish the goal of winning.[9]

9. Regarding the "diminishing sensitivity" result, recall the portion of the value function presented in Chapter 4 defined over

The graph below depicts Berger and Pope's results:



(Berger and Pope 2011)

The upward slope of the hashed line indicates that the more points the home team has at halftime relative to the away team, the more likely the home team will wind up winning the game. The line's discontinuity in the neighborhood of zero depicts the study's main results. At one point behind, the probability of the home team winning is roughly 60%. At one point ahead, the home team's probability of winning drops to roughly 54%, which is slightly higher than if the home team is down by two points at halftime. Similarly, if the home team is ahead by two points at halftime, then its probability of winning is over 60%. Hence, when two teams are within a few points of each other going into halftime, halftime is indeed a game's reference point. And the home team's chances of winning the game diminishes fairly rapidly as it falls further behind going into halftime. This latter result can be taken as evidence that a home team's collective marginal disutility of losing diminishes in concert with its chances of winning. Berger and Pope also find that, all else equal, when the home team is losing at halftime, its probability of winning the game increases by anywhere from 6% to 8%.

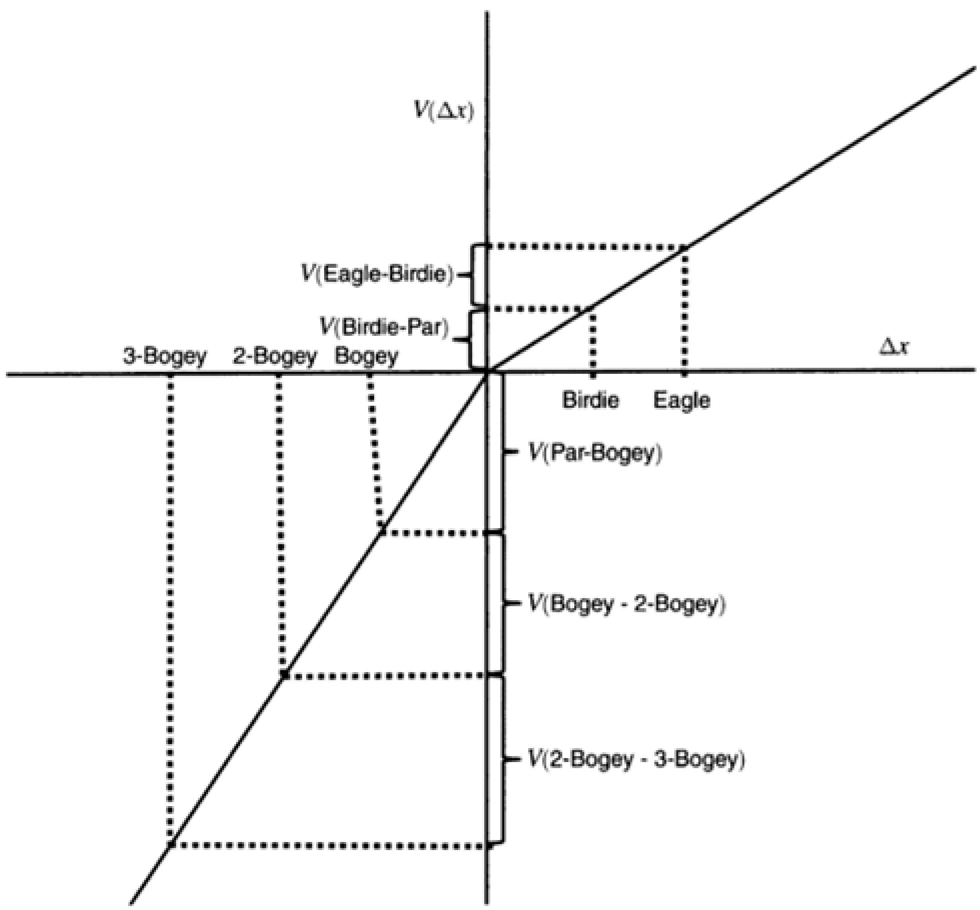
LOSS AVERSION IN PROFESSIONAL GOLF

Professional basketball is not the only sport lending itself to empirical testing of behavioral economics' preeminent theories. Professional golf is also amenable. Using data on over 2.5 million putts measured by laser technology, Pope and Schweitzer (2011) test for the presence of loss aversion among professional golfers competing on the Professional Golf Association (PGA) Tour. As the authors point out, golf provides a natural setting to test for loss aversion because golfers are rewarded for the total number of strokes they take during a tournament, yet each individual hole has a salient reference point, par.

disutility (i.e., losses). Although more steeply sloped than the portion defined over utility (i.e., gains), the portion defined over losses is nevertheless concave shaped.

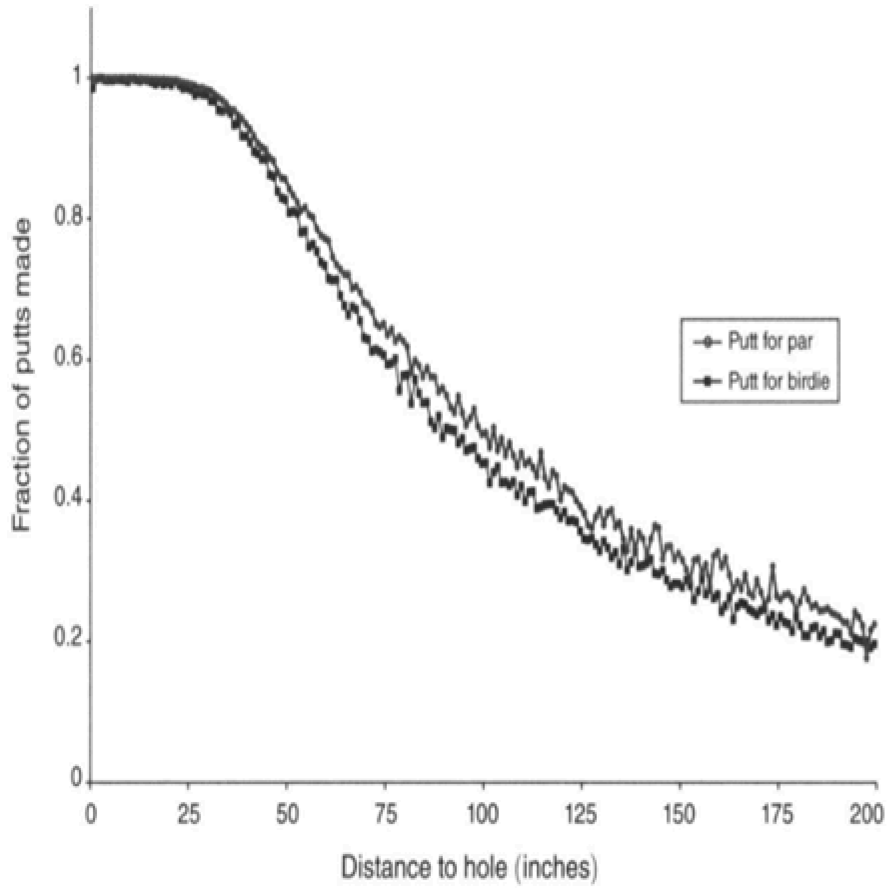
Pope and Schweitzer find that when golfers are "under par" (e.g., putting for a "birdie" that would earn them a score one stroke under par), they are 2% less likely to make the putt than when they are putting for par or are "over par" (e.g., putting for a "bogey" that would earn them one stroke over par). Even the best golfers—including Tiger Woods at the time—show evidence of loss aversion in these situations. Loss aversion motivates golfers to make a higher percentage of puts when they are putting for a bogey than a birdie.

Two figures coalesce the authors' econometric results. The first figure represents the typical golfer's value function. Note the function's reference point (i.e., its origin) at par. The steeper portion of the function defined over the disutility region is associated with missing par and thus bogeying a putt (one-over-par is a bogey, two-over-par is a double bogey, and so on). The flatter portion of the function defined over the utility region corresponds to scoring under par with a birdie (one under par), eagle (two under par), or albatross (greater than two under par). Recall that the relative steepness of the function in the disutility region depicts loss aversion. The linearity of the function indicates an absence of the diminishing effect.



(Pope and Schweitzer 2011)

The next figure depicts the relationships between the average golfer's fraction of putts made when putting for par and for birdie, respectively, relative to distance from the hole. As expected, regardless of whether a golfer is putting for birdie or par, the fraction of putts made decreases as the distance to the hole increases. Of particular interest in this study is that, at each distance, the fraction of putts made is less when the golfer is putting for birdie as opposed to par—again, evidence of loss aversion.



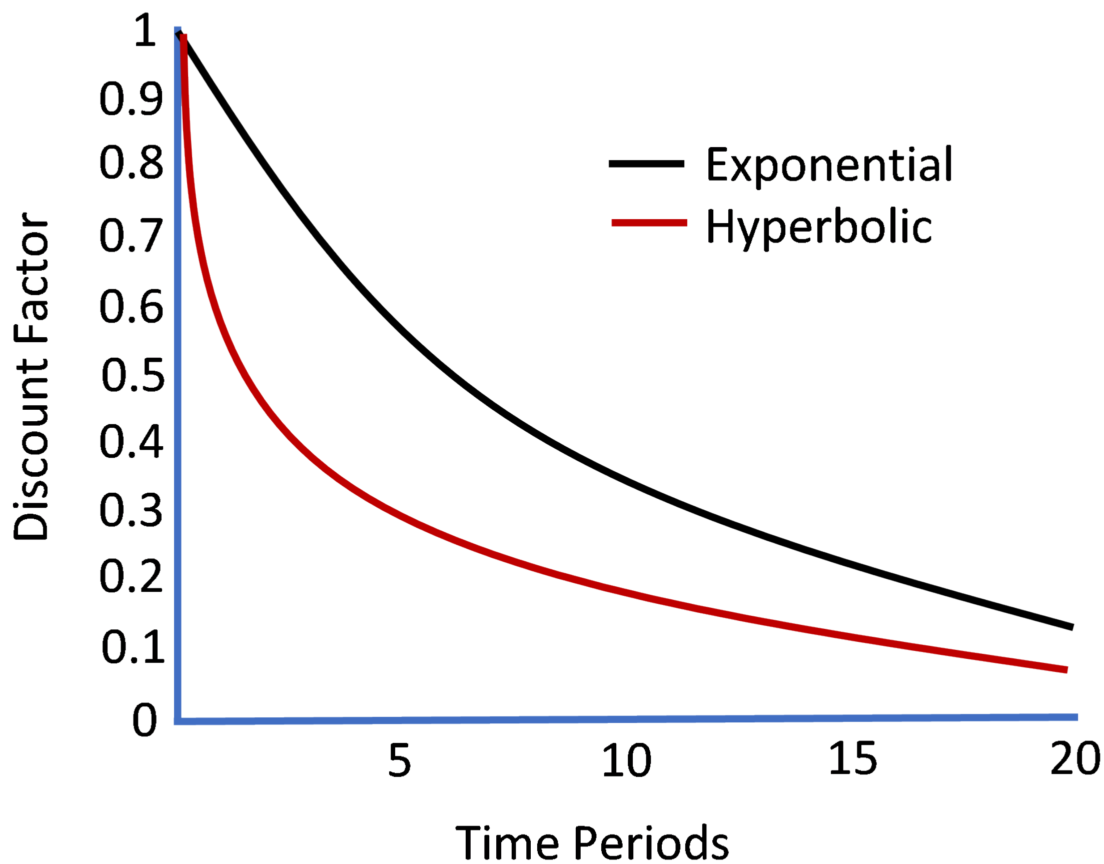
(*Pope and Schweitzer 2011*)

To reiterate, this study's main econometric results reveal a negative effect on sinking a putt when the typical golfer is putting for birdie, and a positive effect on putting for bogey. Consistent with the previous graphs, these numerical results suggest that the typical professional golfer is more likely to sink a put for bogey and less likely to sink the putt for birdie (i.e., the typical golfer is indeed loss averse).[10]

ARE CIGARETTE SMOKERS HYPERBOLIC TIME DISCOUNTERS?

Recall from Chapter 4 the distinction between time-consistent exponential time discounters (*Homo economicus*) and potentially time-inconsistent hyperbolic discounters (*Homo sapiens*). The discounting time paths for exponential versus hyperbolic discounting looked like this:

10. A negative effect associated with putting for double bogey suggests that the typical golfer suppresses his inclination for loss aversion when putting for a score worse than bogey.



A feature distinguishing a hyperbolic from an exponential time discounter is that the former discounts time delays in near-future consumption at much higher rates than the latter, but the discounting of more distant-future consumption converges between the two.

In contrast with Becker and Murphy's (1988) early theoretical work explaining rational addiction among *Homo economicus* based upon exponential time discounting, experimental research aimed at explaining addiction among *Homo sapiens* has found that hyperbolic discounting of future consumption can at least partially explain the impulsive behavior exhibited by those among us with addictions to drugs such as alcohol, heroin, and opioids (c.f., Vuchinich and Simpson, 1998 and Madden et al., 1997).[11] Bickel et al. (1999) also find evidence of hyperbolic time discounting among cigarette smokers. In their field experiment, the authors compare the discounting of hypothetical monetary payments by current and ex-smokers of cigarettes, as well as those who have never smoked (henceforth "never smokers"). For current smokers, the authors also examine discounting behavior associated with delayed hypothetical payment in cigarettes.[12], [13]

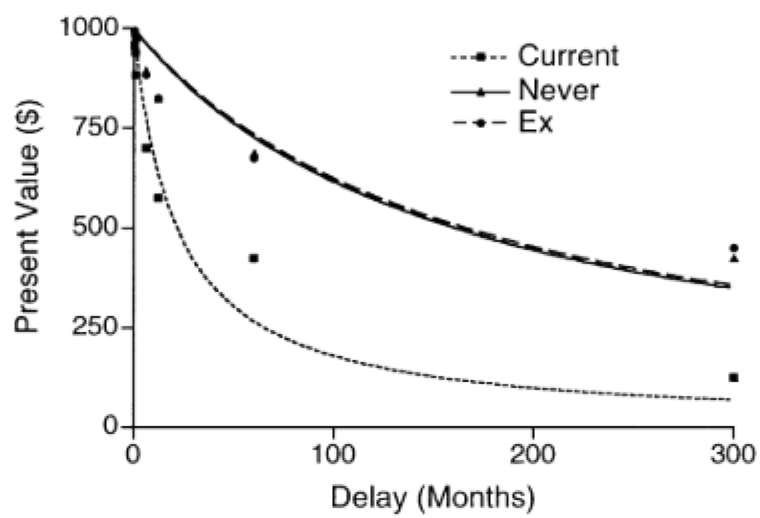
11. Conlin et al. (2007) point out that Projection Bias also manifests itself in the consumption of addictive goods. For example, people may often become addicted to cigarettes, illicit drugs, and alcohol because (1) they underappreciate the negative consequences of being an addict, and (2) they underappreciate how hard it will be to quit once addicted.

12. Current smokers are classified as those who reported smoking at least 20 cigarettes per day for at least the past five years and had a Fagerström Test for Nicotine Dependence (FTND) score of at least six. Never-smokers were those who reported never smoking. Ex-smokers were those who reported abstinence from cigarettes for at least one year, and who had smoked at least 20 cigarettes per day for at least five years prior to quitting.

13. Twenty-six different monetary payment amounts were used to measure the participants' time discounting behavior—$1000, $990, $960, $920, $850, $800, $750, $700, $650, $600, $550, $500, $450, $400, $350, $300, $250, $200, $150, $100, $80, $60, $40, $20, $10, $5, and $1—in concert with seven payment delay periods—one week, two weeks, one

The authors find that current smokers discount the value of a delayed monetary payment more than ex- and never-smokers (the latter two groups do not differ in their discounting behaviors). For current smokers, delayed payment in cigarettes loses subjective value more rapidly than delayed monetary payment. The hyperbolic equation provides a better fit of the data for cigarette smokers than the exponential equation for 74 out of the 89 different comparisons between current cigarette smokers, on the one hand, and ex- and never-smokers on the other. Bickel et al. (1999) conclude that cigarette smoking, like other forms of drug dependence, is characterized by rapid loss of subjective value for delayed outcomes (i.e., pronounced hyperbolic discounting).

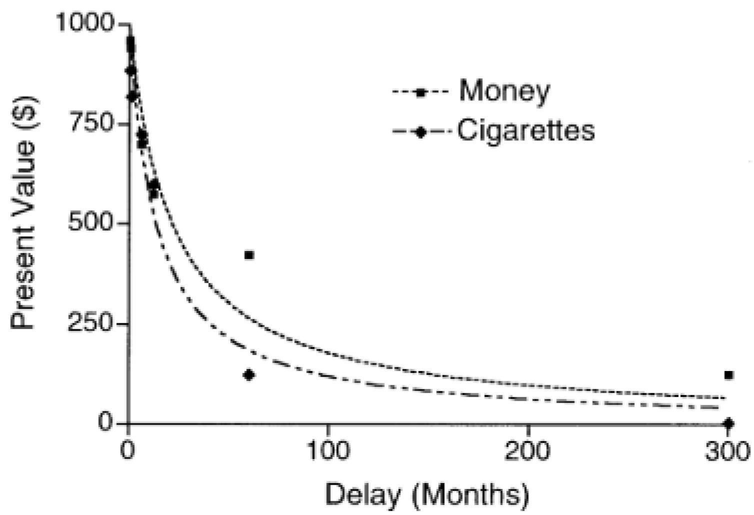
The figure below shows Bickel et al.'s results for a monetary payment scheme. The curves represent the median indifference points (i.e., the estimated values of immediate payment at the respective points of subjective equality with each of seven different delay periods) for current smokers, never-smokers, and ex-smokers. We see that the subjective values decrease more rapidly for smokers (along the curve resembling a hyperbolic discounting function) than for never-smokers and ex-smokers (along curves resembling exponential discounting functions). For example, for smokers, a $1000 payment lost 42.5% of its value when delayed by one year, but for never- and ex-smokers, a $1000 payment lost only 17.5% of its value when delayed by one year.



(Bickel et al. 1999)

As the figure below shows, the bulge in current smokers' hyperbolic discounting function is more pronounced for the cigarette payment scheme (the curve associated with the monetary payment scheme is reproduced from the previous figure for ease of comparison).

month, six months, one year, five years, and 25 years. To ensure that the magnitudes of the monetary and cigarette payment schemes were equal for current smokers, the current smokers were each asked how many cartons of cigarettes they could purchase with $1000. Participants then chose between amounts of money (or cigarettes for the cigarette payment scheme) delivered immediately and corresponding amounts delivered after a given payment delay period.



(Bickel et al. 1999)

ARREST RATES AND CRIME REDUCTION

As Levitt (1998) points out, the linchpin of the rational-choice model of crime is the concept of deterrence: criminal *Homo economicus* will choose to commit fewer criminal acts when faced with higher probabilities of detection or more severe sanctions. Levitt conjectures that criminal *Homo sapiens* may defy this rational-choice model of deterrence by being poorly informed about the likelihood of getting caught, over-optimism about their abilities to evade detection, myopia due to the time gap between committing the crime and imprisonment, or perhaps because serving a prison sentence satisfies a rite of passage among a criminal's peers.

Levitt further points out that empirically testing for a deterrence effect among would-be criminals is fraught with challenges because increasing the expected punishment associated with a given crime can potentially reduce crime through two different channels. The first channel is deterrence—larger penalties and/or higher arrest rates induce criminals to commit fewer crimes. The second channel is incapacitation—if criminals commit multiple offenses and punishment takes the form of imprisonment, increasing expected punishment will also reduce crime by getting criminals off the streets. While a criminal is imprisoned, he is unable to engage in criminal actions that otherwise would have taken place, which biases the statistical effect of deterrence upward (i.e., due to the incapacitation effect, an increase in deterrence measures undertaken by the police would be identified as having a larger negative impact on crime reduction than is truly the case).

Levitt utilized annual reported-crime data from the Federal Bureau of Investigation (FBI) for 59 of the largest U.S. cities over the period 1970-1992 to test for a deterrence effect driven by changes in arrest rates. His results suggest that (1) incapacitation predominately reduces the incidence of rape, (2) incapacitation and deterrence effects are of equal magnitude in reducing the incidence of robbery, and (3) the deterrence effect outweighs the incapacitation effect in reducing aggravated assault and property crimes (Levitt estimates that the deterrence effect accounts for more than 75% of the observed effect of arrest rates on property crime).

Hence, when it comes to arrest rates, criminal *Homo economicus* and *Homo sapiens* share similar responses to deterrence.

INTERPERSONAL DYNAMICS IN A SIMULATED PRISON

There is substantial evidence that prisons in the US (if not worldwide) neither rehabilitate prisoners nor deter future crime. In its most recent report on recidivism in the US, the US Justice Department reports that 44% of state prisoners released in 2005 across 30 states were re-arrested within one year of their release, 68% within four years, 79% within six years, and 83% within nine years (Alper et al., 2018). Of released drug offenders, 77% were re-arrested for a non-drug crime within nine years after release. During each year, and cumulatively during the nine-year follow-up period, released non-violent offenders were more likely than released violent offenders to be arrested again (Alper et al., 2018).

Haney et al. (1973) pose (and then seek to answer) a nagging question pertaining to what lies behind these statistics. To what extent can the deplorable conditions of our penal system and their often-dehumanizing effects upon prisoners and guards—conditions that likely contribute to recidivism—be explained by the nature of the people who administer it (prison guards) and the nature of the people who populate it (prisoners)? The authors' dispositional hypothesis is that a major contributing cause of these conditions can indeed be traced to some innate or acquired characteristics of the correctional and inmate populations. As the authors point out, the hypothesis has been embraced by both the proponents of the prison status quo, who blame the nature of prisoners for these conditions, as well as the status quo's critics, who blame the motives and personality structures of guards and staff.

To understand the genesis of prison culture—in particular, the cultural effect on the disposition of both prisoners and guards—Haney et al. (1973) undertook one of the most notorious (or, depending upon one's perspective, noteworthy) field experiments ever conducted with willing, non-incarcerated adults. The authors designed a functional simulation of a US prison in which subjects who were drawn from a homogeneous, "normal" sample of male college students role-played prisoners and guards for an extended period of time. Half the subjects were randomly assigned to the prisoner group, which was incarcerated for nearly one full week. The other half were randomly assigned to the prison guard group, which played its role for eight hours each day. The behaviors of both groups were observed, recorded, and analyzed by the authors, particularly regarding transactions occurring between and within each group of subjects.

The 21 subjects who ultimately participated in the experiment (out of a total of 75 applicants) were judged to be the most physically and emotionally stable, most mature, and exhibited the least anti-social behavior. The prison was constructed in a basement corridor in the Psychology Department's building at Stanford University. It consisted of three small cells (6' x 9'), each cell housing three prisoners. A cot, mattress, sheet, and pillow for each prisoner were the only pieces of furniture in each cell. A small, unlit closet across from the cells (2' x 2' x 7') served as a solitary confinement facility. Several rooms in an adjacent facility were used as guards' rooms and quarters for a "warden" and "superintendent." The prisoners were each issued identical, ill-fitting, prisoner uniforms to instill uniformity and anonymity in the prisoners' daily existence. The guards' uniforms consisted of a plain khaki shirt and trousers, a whistle, wooden baton, and reflecting sunglasses that made eye contact impossible.

With help from the Palo Alto City Police Department, the prisoners were each "arrested" (with handcuffs, no less) at their residences under suspicion of burglary and armed robbery, taken to the police station, and "processed" under normal induction procedures. Once they arrived at the simulated prison (blindfolded, no less), they continued with standard induction procedures, which included being stripped naked, sprayed with a deodorant, and made to stand alone naked in a prison yard for a short period of time. Each prisoner was then put in his cell and ordered to remain silent. During their confinement, the prisoners were fed three meals a day, allowed three supervised toilet visits, and were allotted two hours daily for the privilege of reading and letter-writing.

Data was gathered via videotaping, audio recordings, personal observations, and a variety of checklists filled out by the guards and researchers. Through subsequent analysis of the data, Haney et al. (1973) found that the personal behaviors of the prisoners and guards, and the social interactions between them, supported many commonly held conceptions of prison life and validated anecdotal evidence provided by real-life ex-convicts. In general, both prisoners and guards tended toward increased negativity over the week in terms of their dispositions. For both prisoners and guards, self-evaluations became more disapproving as their experiences were internalized. Prisoners generally adopted a passive response mode while guards assumed active, initiating roles in all prisoner-guard interactions.[14]

Specifically, Haney et al. found that the extent to which a prisoner scored high on his personality test for rigidity, adherence to conventional values, and acceptance of authority helped determine the likelihood that he adjusted more effectively to the authoritarian prison environment. In written self-reports, prisoners expressed nearly three times as much negativity as positivity. Guards expressed slightly more negativity than positivity. Prisoners also showed roughly three times as much mood fluctuation as did the guards.

Haney et al. conclude:

> *"The conferring of differential power on the status of "guard" and "prisoner" constituted, in effect, the institutional validation of those roles. But further, many of the subjects ceased distinguishing between prison role and their prior self-identities. When this occurred, within what was a surprisingly short period of time, we witnessed a sample of normal, healthy American college students fractionate into a group of prison guards who seemed to derive pleasure from insulting, threatening, humiliating, and dehumanizing their peers—those who by chance selection had been assigned to the prisoner role. The typical prisoner syndrome was one of passivity, dependency, depression, helplessness and self-deprecation." (p. 89)*

For those of us who are skeptical of the simulated nature of this experiment's constructed prison environment, Haney et al. offer this final thought:

> *"In one sense, the profound psychological effects we observed under relatively minimal prison-like conditions which existed in our mock prison make the results even more significant and force us to wonder about the devastating impact of chronic incarceration in real prisons." (p. 91)*

At the very least, this experiment demonstrates how manipulatable and culpable *Homo sapiens* can become in the context of a field experiment.[15]

CORRUPTION IN SUMO WRESTLING

In one of their most well-known studies, Duggan and Levitt (2002) uncovered the extent of

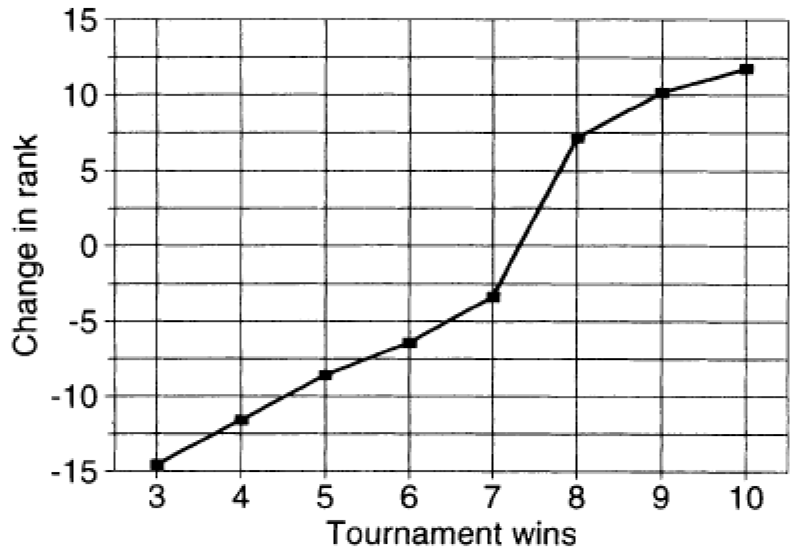
14. As Haney et al. report, "the most dramatic evidence of the impact of this situation upon the participants was seen in the gross reactions of five prisoners who had to be released because of extreme emotional depression, crying, rage, and acute anxiety" (Page 81). Today, the need for what is known as institutional review board (IRB) pre-approval of human subjects research makes research of this nature impermissible.

15. In what may be the most well-known field experiment designed to induce obedience to authority, Milgram's (1963) participants were led to believe that they were assisting in an unrelated experiment in which they were instructed to administer electric shocks to an unseen "learner." The participants gradually increased the levels of the electric shocks (which, unbeknownst to them, were fake) to levels that would have been fatal had they instead been real.

corruption in Japan's national sport, sumo wrestling. To understand how they did so, one must know something about how sumo wrestling tournaments work.

A sumo tournament involves 66 wrestlers (*rikishi*) competing in 15 bouts each. A wrestler who achieves a winning record (eight wins or more) in a tournament is guaranteed to rise in the official ranking of the nation's wrestlers. A wrestler with a losing record in the tournament (seven wins or less) falls in the national rankings. A wrestler's ranking is a source of prestige and the basis for salary determination and various in-kind perks.

As Duggan and Levitt point out, the key institutional feature of sumo wrestling making it ripe for corruption is the concomitant nonlinearity in the ranking (and thus payoff) function for competitors, depicted in the figure below:

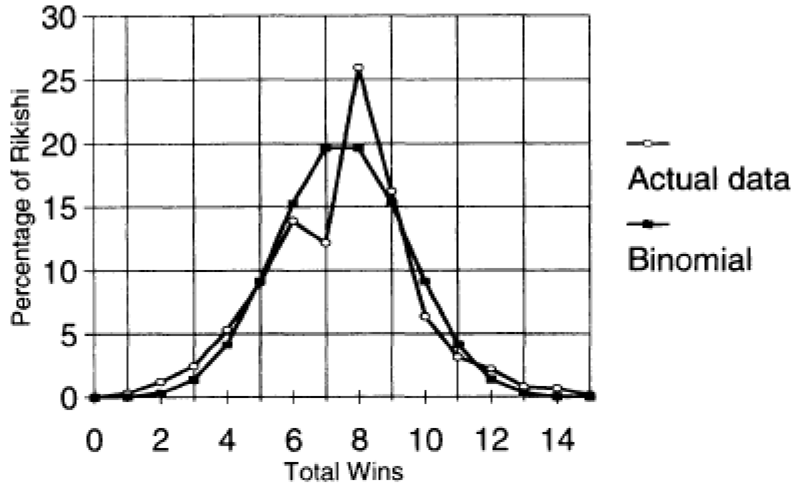


(Duggan and Levitt 2002)

We see that a wrestler who achieves a losing record of seven wins and eight losses (7-8) can expect to drop in the rankings by roughly three places (e.g., if, going into the tournament, the wrestler was ranked third nationally, after the tournament, he is now ranked sixth). To the contrary, a wrestler achieving a winning record of 8-7 in the tournament can expect to rise in rank by roughly eight places. Consequently, a wrestler entering the final match of a tournament with a 7-7 record has far more to gain from a victory than an opponent with a record of, say, 8-6 has to lose.

Following almost 300 wrestlers from 1989-2000, the authors find that wrestlers who are on the margin for attaining their eighth victory in a given tournament (in what's known as a "bubble match") win far more often than one would expect. Further, whereas the wrestler who is on the margin for his eighth victory in a bubble match wins with a surprisingly high frequency, the next time the same two wrestlers face each other in another tournament, it is the opponent (i.e., the wrestler who threw the bubble match) who has an unusually high win percentage. In other words, Duggan and Levitt not only uncover corruption in the bubble match itself but also corruption in the subsequent match between the same two wrestlers. This corruption comes in the form of the earlier bubble match's winner duly compensating the loser by similarly throwing the current match.

Duggan and Levitt depict their finding in the figure below:



(Duggan and Levitt 2002)

The figure shows two curves—one based on the actual data, the other based on the binomial distribution, which represents the distribution we would expect to hold between the wrestlers and their wins, all else equal. The binomial distribution depicts a nice, bell-shaped curve where the largest percentages of wrestlers win between 5 and 10 matches per tournament. The obvious spike in the actual data over eight wins, which is aligned with over 25% of the wrestlers when we would expect only 20%, suggests a preponderance of unexpected outcomes in bubble matches.

Interestingly, Duggan and Levitt find that the bubble match effect disappears in tournaments with high levels of media scrutiny and when the opponent (i.e., the wrestler who would otherwise agree to throw the bubble match) is in the running for one of the tournament's special prizes.[16] By contrast, success on the bubble increases for veteran wrestlers (i.e., all else equal, veterans are more likely to win bubble matches in tournaments where they go into the match with seven wins, seven losses).

CORRUPTION IN EMERGENCY AMBULANCE SERVICES

To improve emergency ambulance response times in England in the early 2000's, authorities implemented a common response-time target for "ambulance trusts" (i.e., regional units) that 75% of potentially immediately life-threatening (Category A) emergency telephone calls be met within 8 minutes of the call having been placed. Less serious emergency calls (e.g., concerning serious but not life-threatening or neither serious nor life-threatening) were assigned less stringent targets. In addition, a "star rating system" was established rewarding or penalizing the trusts based upon the extent to which they met or did not meet the targets.

As Bevan and Hamblin (2009) point out, hospital rankings based upon the annual star ratings were easy to understand, and the results were widely disseminated (published in national and local newspapers and on websites, and featured on national and local television). Hospital staff was highly engaged with the information used to determine the ratings. Further, the star ratings mattered for chief executives, as being zero-rated resulted in damage to their professional reputations and affected staff recruitment. As a result, the star rating system was widely considered to be a salient mechanism

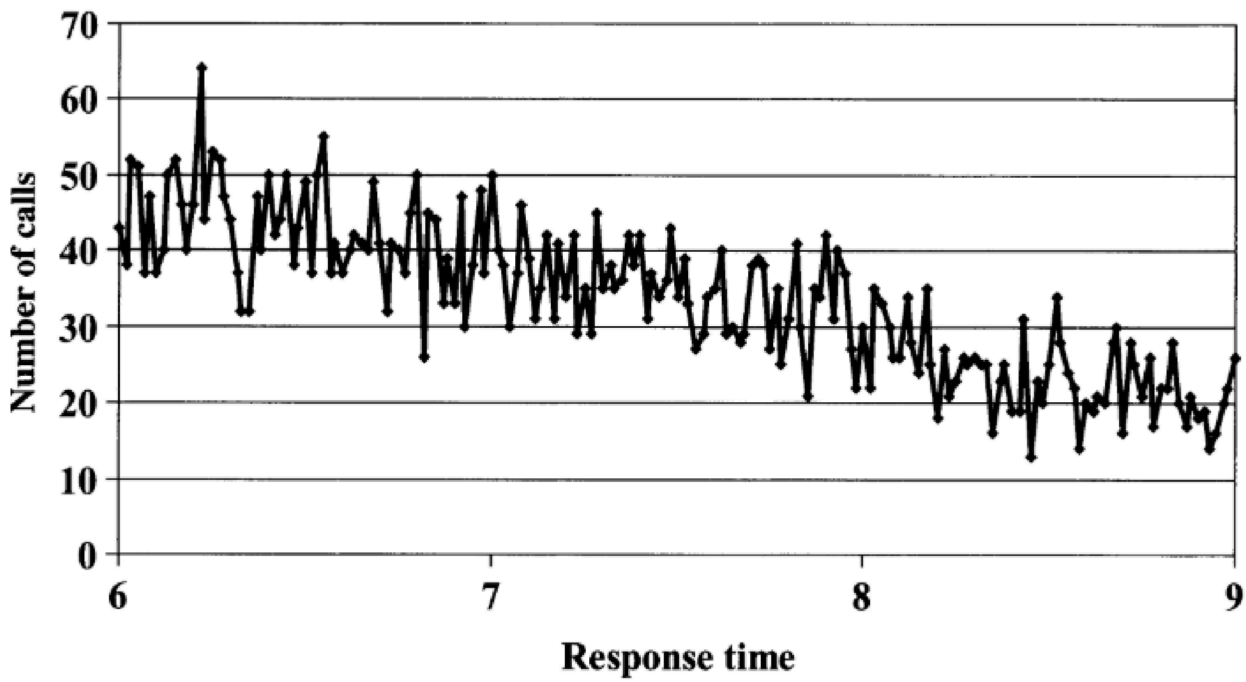
16. With respect to the media's effect on corruption in bubble matches, recall Pope et al.'s (2018) similar result for racial discrimination in the NBA.

for improving hospital performance and, as a result, was ripe for attempts by hospitals and ambulance trusts to manipulate it.

Bevan and Hamblin find that, on the surface, the implementation of Category A ambulance-service targets in 2002 had a noticeable impact on response times. The percentage of response times per trust meeting the eight-minute target increased markedly after 2002 and remained up in the range of 70%–90% meeting the target through the study period of 2005.

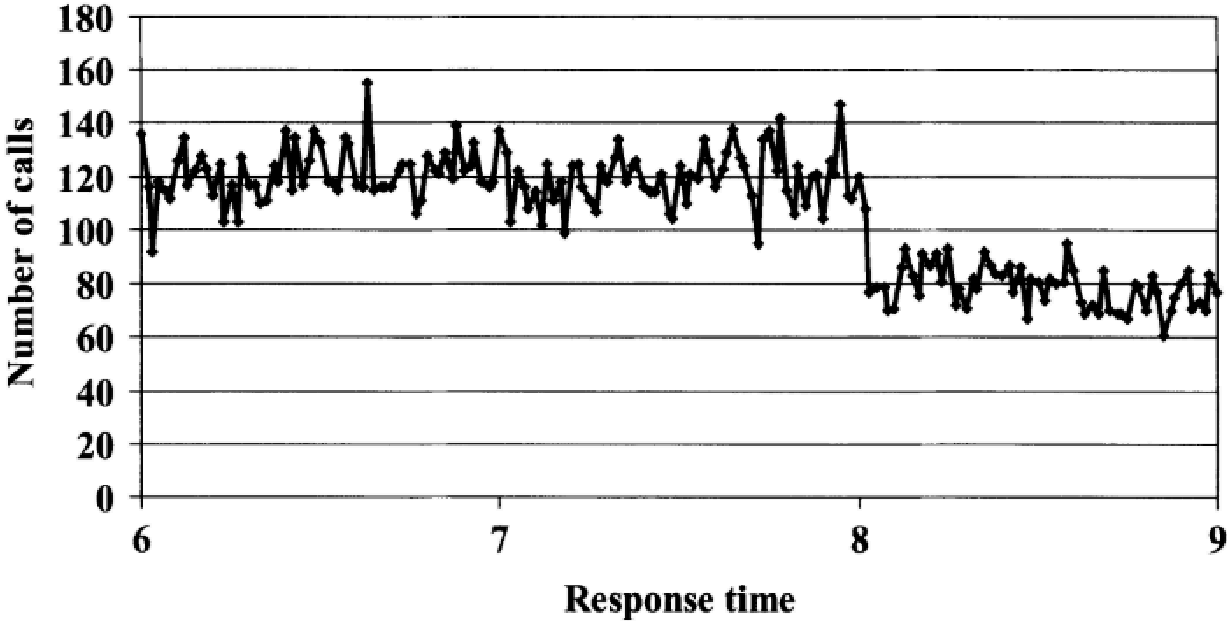
However, digging deeper into the data, Bevan and Hamblin uncovered pervasive evidence of cheating among the trusts. As the authors point out, the system's intense focus on the Category A target gave rise to several concerns, among them the obvious incentive to classify calls as Categories B and C rather than Category A, and the fact that arriving at the scene in 8.01 minutes was now inevitably seen as a failure. Earlier investigations had concluded that the former concern—reduced number of calls classified as Category A—was not commonly practiced among the trusts. Not so the latter concern. Bevan and Hamblin find that among the trusts' response times taking longer than the targeted eight minutes, roughly 30% had been 'corrected', i.e. re-recorded as having taken less than eight minutes.

First, consider the recorded response-time data for a trust that exhibited an expected ('uncorrected') distribution of response times—a "noisy" decline in the number of responses with no obvious jump around the eight-minute threshold:

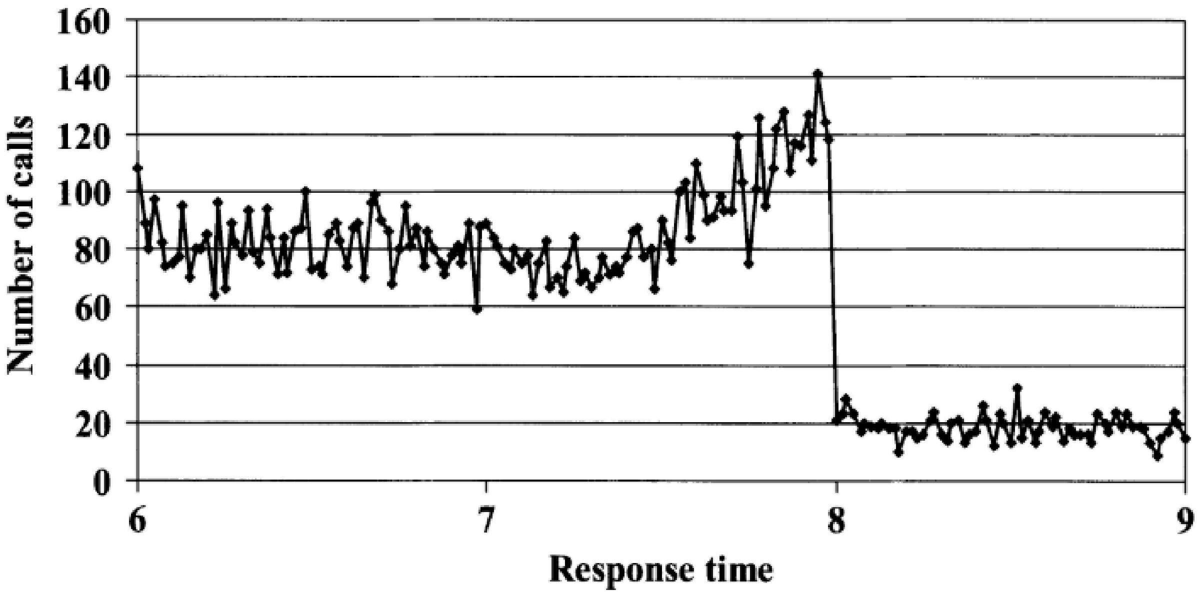


(Bevan and Hamblin 2009)

Next, consider data from two other trusts that exhibit what appear to be curious drops in reported response times at the 8-minute threshold:



(Bevan and Hamblin 2009)



(Bevan and Hamblin 2009)

The drop in reported response times is obviously more marked in the bottom figure, but also present in the first of these two figures. Clearly, something suspicious occurred with the reporting for these two trusts. As with sumo wrestlers, the putative setting of a harmless rule induced perverse behavior among the targeted group of *Homo sapiens*. In the case of England's emergency ambulance services, it appears that some of the ambulance trusts chose to disingenuously fudge their reported Category A response times.

NEW YORK CITY'S TAXI CAB DRIVERS

Camerer et al. (1997) clued into the fact that taxi cab drivers are an ideal population to study for unexpected labor market behavior because the structure of the taxi cab market (at least, New York City's (NYC's) market in the late 1980s and early 1990s) enabled drivers to choose how many hours to drive during a given shift. As a result, drivers faced wages that fluctuated daily due to "demand shocks" caused by weather, subway breakdowns, day-of-the-week effects (e.g., Mondays may generally be busier than Tuesdays each week), holidays, conventions, etc. Although rates per mile are set by law, on busy days, drivers may have spent less time searching for customers and thus, all else equal, earned a higher hourly wage. These hourly wages are transitory. They tend to be correlated within a given day and uncorrelated across different days. In other words, if today is a busy day for a driver, she can earn a relatively high hourly wage. But if the very next day is slow, then the driver will earn a relatively low hourly wage.

Camerer et al. compiled different samples of NYC taxi drivers over three different time periods: (1) from October 29th to November 5th, 1990, consisting of over 1000 trip sheets filled out by roughly 500 different drivers (henceforth the TLC1 sample), (2) from November 1st to November 3rd, 1988, consisting of over 700 trip sheets filled out by the same number of drivers (henceforth the TLC2 sample), and (3) during the spring of 1994, consisting of roughly 70 trip sheets filled out by 13 different drivers (henceforth the TRIP sample). For each sample, Camerer et al. divided drivers into low- and high-experience subsamples.

Generally speaking, the authors find that drivers (particularly inexperienced ones) made labor supply decisions "one day at a time" (i.e., framed narrowly) rather than inter-temporally substituting their labor and leisure hours across multiple days (i.e., framed broadly) in response to temporary hourly wage changes (as you've probably guessed already, *Homo economicus* drivers frame broadly). The typical (*Homo sapiens*) driver set a loose daily income target (which served as the driver's reference point) and quit working once she reached that target (which resulted in a negative relationship between the number of hours she chose to work and the driver's daily hourly wage rate). In other words, as the driver's hourly wage rose, she chose to drive fewer hours—a perverse outcome in a rational-choice model of any type of worker's behavior. As Camerer et al. point out, the driver's reference point established a daily mental account and also suggests loss-averting behavior in the sense that, on a slow day, a driver chose to work more hours to reach the reference point, thus avoiding the "loss" that comes with under-performing on the job.

Specifically, the authors find that low-experienced drivers exhibit negative responses to wage increases in each sample, but the responses are statistically significant only in the TRIP sample and marginally significant in the TLC2 sample. High-experienced drivers exhibit a negative response solely in the TLC1 sample. Therefore, Camerer et al. find some evidence of reference dependency, mental accounting, and loss aversion among NYC's famed taxi drivers.

SAVINGS PLANS FOR THE TIME-INCONSISTENT

Homo sapiens who are time inconsistent when it comes to saving income for future consumption are prone to save too little now for what they later realize they needed in order to maintain their standard of living. In response, two types of "tailored savings plans" have been developed over time, targeting segments of the population with historically low personal savings rates. One plan—Prize-Linked Savings Accounts (PLSAs)—encourages people to increase their savings rates by adding a lottery component to what is an otherwise traditional savings account at a participating bank (Morton, 2015). Depositors' accounts are automatically entered into periodic drawings based upon their account balances during a given period. Depositors then have a chance to win prizes, which are funded through the interest that accrues across the pool of PLSAs held at the bank.

As Morton points out, although they are relatively new in the US, PLSAs have a long history internationally. The first known program was created in the United Kingdom (UK) in 1694 as a way to pay off war debt. PLSAs are currently offered in 22 countries, including Germany, Indonesia, Japan, and Sweden. Because of Americans' relatively low personal savings rates, and, as pointed out in Section 1, *Homo sapiens*' general propensity to overweight improbable events (and thus, to accept gambles), PLSAs could potentially help raise savings rates in the US.

The US personal savings rate hit a high of 17% of disposable personal income in 1975, declining to roughly 2% by 2005, before rebounding to roughly 5% by 2014 (Morton, 2015). An estimated 60% of Americans had less than $1,000 in personal savings in 2018 (Huddleston, 2019). And yet, in 2019 an estimated 44% of American adults visited a casino (American Gaming Association, 2019). Hence, it seems that statistics also point to the potential role that PLSAs can play in nudging *Homo sapiens* to save more of their personal income.

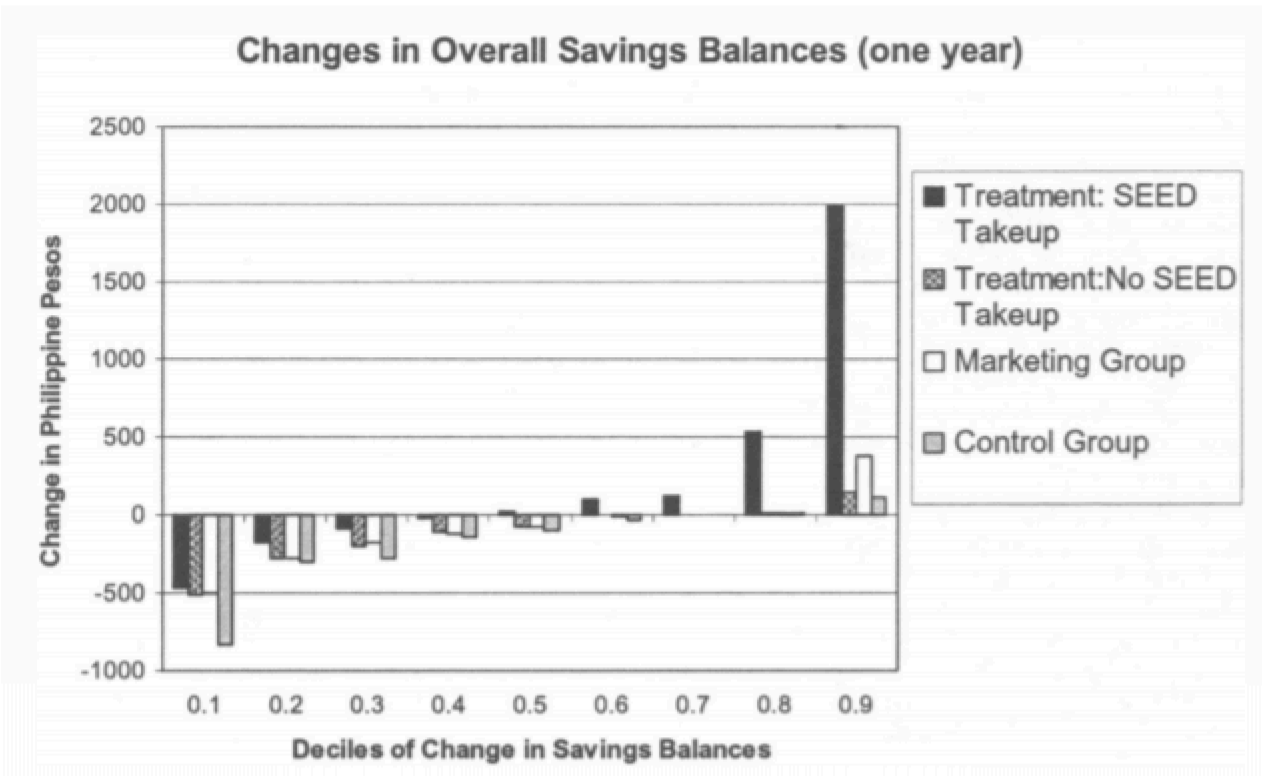
One motivation behind the establishment of PLSAs is that *Homo sapiens* suffer from time-inconsistency when it comes to committing to saving for their futures. For some prospective savers, this time-inconsistency problem manifests itself as procrastination in opening up a savings account. For others, saving for the future is not considered imperative when juxtaposed against the need to cover current expenses.

A second type of tailored savings plan—Commitment Savings Accounts (CSAs)—involves a prospective saver, or client, specifying a personal savings goal upfront which can be either date-based (e.g., saving for a birthday or wedding) or amount-based (e.g., saving for a new roof). The client decides for himself what the goal will be and the extent to which his access to the account's deposits will be restricted until the goal is reached. The CSA earns the same rate of interest as a normal bank account.

To test the efficacy of a CSA in helping clients overcome their time-inconsistent savings decisions, Ashraf et al. (2006) conducted a field experiment with over 1,700 existing and former clients of Green Bank of Caraga, a rural bank in the Philippines. The authors first conducted a survey of each client to determine the extent of his or her time-inconsistency problem (i.e., to determine whether the client is an exponential time discounter (which, as we learned in Chapter 3, describes *Homo economicus*), a hyperbolic time discounter (which, as we learned in Chapter 4, describes many a *Homo sapiens*), or perhaps an inverted hyperbolic time discounter whose discount rate actually rises as the time delay for receiving a reward increases (recall that, under hyperbolic discounting, this rate falls as the time delay increases)). Next, half of 1,700 clients were randomly offered the opportunity to open a CSA, called a SEED account in this particular instance (Save, Earn, Enjoy Deposits)—the study's treatment group. Of the remaining half of clients, half received no further contact (the study's control group) and half were encouraged to save at a higher rate using one of the bank's more traditional accounts (the study's "marketing group").

Of the subsample of clients in the treatment group, roughly 28% chose to open SEED accounts with the bank, the majority of which were date-based. After 12 months, just under 60% of the SEED accounts reached maturity (if date-based) or reached the threshold amount (if amount-based), and all but one client chose to open a new SEED account thereafter. Also, account balances for SEED account holders were markedly higher than for those clients in both the marketing and control groups. Further, women identified as hyperbolic discounters prone to time-inconsistent savings behavior (and thus, who presumably have stronger preferences for the SEED account's commitment mechanism) were significantly more likely to open a SEED account. Preferences for the SEED account among time-inconsistent men were not as strong.

The figure below provides evidence of the SEED account's effectiveness in inducing higher savings balances among those clients in the experiment's treatment group who chose to open an account. Compared with clients in the control and marketing groups, as well as those in the treatment group who chose not to open a SEED account (Treatment: No SEED take-up), clients in the treatment group who opened a SEED account (Treatment: SEED take-up) grew larger savings balances after one year, especially among those clients with the largest balances (i.e., from the 0.6 to 0.9 decile groupings). Among those clients who suffered losses in their savings balances by year's end, the losses suffered by the Treatment: SEED take-up clients were the smallest (as depicted for the 0.1 to 0.5 decile groupings).



(*Ashraf et al. 2006*)

As the results of this study suggest, tailored savings plans such as SEED appear to have potential for taking the "in" out of *Homo sapiens'* time-"in"consistent tendencies when it comes to saving for the future.

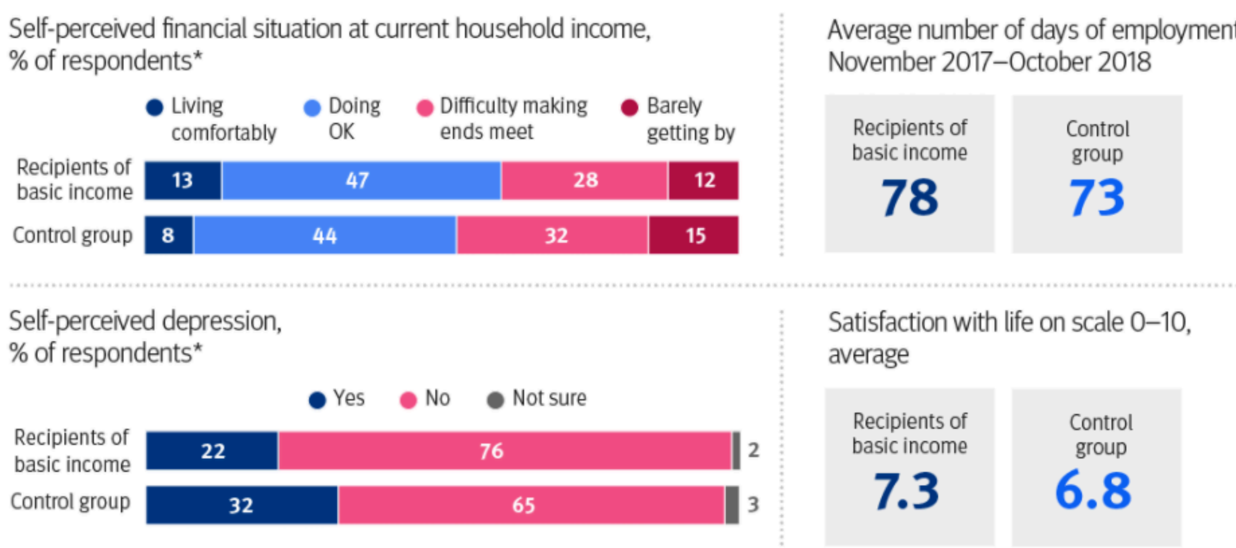
THE FINNISH BASIC INCOME EXPERIMENT

Most nations provide some form of public social expenditure (PSE) to assist lower-income and otherwise marginalized citizens in meeting their basic needs over time. For example, among Organization for Economic Cooperation and Development (OECD) countries, the nations of France, Belgium, Finland, Denmark, Italy, Austria, Sweden, Germany, and Norway devote at least 25% of their

Gross Domestic Products (GDPs) to PSE (OECD, 2019). PSE includes cash benefits, expenditures on health and social services, public pension payments, and unemployment and incapacity benefits.

In 2017, the Finnish government conducted a two-year field experiment to learn if providing a basic income in lieu of PSE might boost employment and well-being among recipients more effectively than its traditional PSE programs (Kangas et al., 2019). In the experiment, a treatment group of 2,000 randomly selected unemployed persons between the ages of 25 and 58 received a monthly payment of €560 unconditionally and without means testing. The €560 monthly payment corresponded to the monthly net amount of the basic unemployment allowance and labor-market subsidy provided by Kela (the Social Insurance Institution of Finland). To study the effects of this basic-income program, the employment and well-being impacts experienced by the treatment group were compared against a control group comprised of 173,000 individuals who were not selected to participate in the experiment.

As the figure below shows, results for the first year of the program indicate that members of the treatment group on average experienced a (statistically insignificant) five-day increase in employment relative to members of the control group (Kela, 2020). Further, on a 10-point life-satisfaction scale, treatment group members reported a (statistically significant) 0.5-point gain.



(*Kela 2020*)

As Kela (2020) points out, although the employment increase was relatively small overall, for families with children who received a basic income, employment rates improved more significantly during both years of the experiment. In general, members of the treatment group were more satisfied with their lives and experienced less mental strain, depression, sadness, and loneliness. They also reported a more positive perception of their cognitive abilities (i.e. memory, learning, and ability to concentrate), and perceived their financial situations as being more manageable.

These results beg an important question when it comes to implementation of new and innovative PSE programs: In the absence of tangible results, such as changes in employment rates, are the intangible benefits experienced by participating *Homo sapiens* worth the social investment?

MICROFINANCE

One of the more innovative approaches to financing small businesses in lower-income countries is known as microfinance (Banerjee, 2013; Mia et al., 2017). Bangladeshi social entrepreneur and 2006 Nobel Prize winner Mohammad Yunus is credited as being the progenitor of microfinance because of a project he initiated in 1976, providing small business loans to small groups of poor residents in rural Bangladeshi villages. The project subsequently led to the founding of Grameen Bank in 1983, whose guiding principle is that small, well-targeted loans are better at alleviating poverty than donor aid.

The basic idea behind microfinance is simple. Because traditional lending requirements in the banking industry rely on borrowers pledging significant collateral to protect the interests of the lender, and because the risk of the borrower defaulting on a bank loan is often large and potentially costly, bank loans are generally considered off-limits to poorer entrepreneurs. Microfinance solves this loan-inaccessibility problem by lending to groups of entrepreneurs who essentially form cooperatives to advance collective business interests and take collective responsibility for loan repayment. The pooling of risk within the group lowers the chance of default on a loan and helps ensure that the loan will be profitable for both the borrower and the lender—a classic "win-win" solution, at least for *Homo economicus* borrowers and lenders.

But what about *Homo sapiens*? Although evidence suggests that microfinance has typically been a win for *Homo sapiens* lenders in terms of high rates of loan repayment (and therefore, low default rates) (Banerjee, 2013; Mia et al., 2017), the proverbial jury is still out regarding the extent to which microfinance has been a win for *Homo sapiens* borrowers. In an extensive field experiment, Banerjee et al. (2015) surveyed a large sample of residents located in 50 randomly selected poor neighborhoods in Hyderabad, India where branches of the microfinance firm Spandana and, later, other firms, had recently been established.[17] The authors surveyed the members of their sample three separate times—in 2005, 2007, and 2009 (i.e., before, during, and after the opening of the Spandana branches).[18]

The authors found that borrowers used microfinance loans to purchase durable goods for their new or existing businesses that had hitherto been unaffordable without the loan money. The typical borrower repaid the loan by reducing consumption of everyday "temptation goods" and working longer hours. No evidence was found of the loans ultimately helping to lift borrowers out of poverty in terms of improved health, education, and empowerment. If the loans helped anyone, it was the relatively larger, already-established businesses with relatively high pre-existing profit levels. Less than 40% of eligible, or "likely borrowers" availed themselves of the microfinance loans even though they continued to borrow from other informal sources.

The evidence for micro-financed loans on the profitability of solely new businesses is likewise bleak. The authors find that new businesses between roughly the 35th and 65th percentiles of profitability have statistically significant lower profits in the neighborhoods where microfinance loans became available. Nevertheless, Banerjee et al. (2015) report that this overall result shields divergent effects across industry types. In particular, new food businesses (tea/coffee stands, food vendors, small grocery stores, and small agriculture) that availed themselves of micro-financed loans on average experienced an 8.5% bump in profitability relative to new food businesses that established themselves in neighborhoods without access to microfinance loans. In contrast, new rickshaw/driving businesses backed by microfinance loans experienced a 5.4% decline in profitability relative

17. Because the criteria for loan eligibility were that a potential borrower be a woman, aged 18-59, and having resided in the area for at least one year, the study's sample consisted solely of individuals who met these criteria.

18. For an interesting laboratory experiment addressing the propensity of Homo sapiens to underreport their earnings from their micro-financed business in order to reduce the level at which they would otherwise be considered capable of repaying the loan, see Abbink et al. (2006).

to new rickshaw/driving businesses that established themselves in neighborhoods without access to microfinance loans.

In conclusion, Banerjee et al. are balanced in their assessment of the findings. They conclude that microfinance is indeed associated with some business creation—in the first year after obtaining microfinance, more new businesses are created, particularly by women. However, these marginally profitable businesses are generally smaller and less profitable than the average business in the neighborhood. Microfinance also leads to greater investment in existing businesses and an improvement in the profitability of the most profitable among those businesses. For other businesses, profits do not increase, and, on average, microfinance does not help these businesses expand in any significant way. Even after three years of having assumed a microfinance loan, there is no increase in the number of these businesses' employees (i.e., business size) relative to businesses that did not assume loans.

Once again, the fickleness of *Homo sapiens* plays itself out in a market setting, this time in the neighborhoods of Hyderabad, India.

TRUST AS SOCIAL CAPITAL

In Section 2 we investigated the trust game and the extent to which *Homo sapiens* participating in laboratory experiments express both their trust and trustworthiness. Knack and Keefer (1997) seek to answer the question, do societies comprised of more trusting and trustworthy individuals, all else equal, perform better on a macroeconomic scale? What is the relationship between interpersonal trust and norms of civic cooperation (i.e., social capital) on the one hand, and economic performance on the other?[19]

As the authors point out, conventional wisdom suggests that economic activities requiring agents to rely upon the future actions of others (e.g., transactions involving goods and services that are provided in exchange for future payment; employment contracts in which managers rely on employees to accomplish tasks that are difficult to monitor; or investments and savings decisions that rely on assurances by governmental agencies or banks that assets will not be appropriated) are accomplished at lower cost in higher-trust societies. Individuals in higher-trust societies spend less time and money protecting themselves from being exploited in economic transactions. Written contracts are less likely to be needed, and when needed, they are not required to specify every possible contingency. Litigation may be less frequent. Individuals in high-trust societies are also likely to divert fewer resources to protecting themselves from unlawful violations of their property rights (e.g., through bribes or private-security services and equipment). Further, high trust can encourage innovation. If entrepreneurs are required to devote less time to monitoring possible malfeasance committed by partners, employees, and suppliers, then they have more time to devote to innovation in new products or processes.

For their measures of trust and civic norms, Knack and Keefer utilize The World Values Survey, which contains survey data on thousands of respondents from roughly 30 different market economies worldwide. The survey question used to assess the level of trust in a society is this:

> *"Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?"*

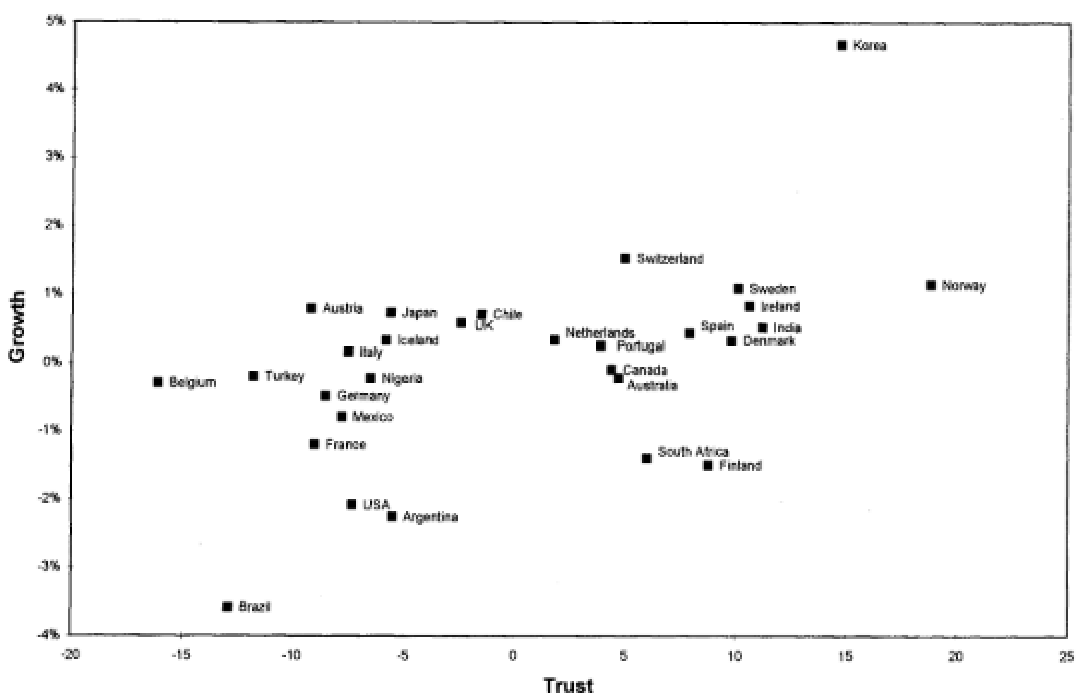
19. Below we look at studies that have explored the role of descriptive and injunctive norms and their saliency in reducing the ill effects of such social predicaments as littering, environmental theft, and drunk driving, as well as encouraging social enhancements such as energy conservation.

Based upon survey participants' responses, the authors created a trust indicator variable (*TRUST*) equal to the percentage of respondents in each nation replying that most people can be trusted. The extent of civic norms present in a given society is gleaned from responses to questions about whether each of the following behaviors can always be justified, never be justified, or something in between:

1. *"claiming government benefits which you are not entitled to"*
2. *"avoiding paying a fare on public transport"*
3. *"cheating on taxes if you have the chance"*
4. *"keeping money that you have found"*
5. *"failing to report damage you've done accidentally to a parked vehicle"*

Respondents chose a number from one (never justifiable) to 10 (always justifiable). The authors summed values over the five items to create a scale (*CIVIC*) with a 50-point maximum score. They then measured the impact of *TRUST* and *CIVIC* on both national growth (in terms of Gross Domestic Product (GDP)) and investment rates. To control for other determinants found in the literature on economic growth, Knack and Keefer included in their regression analysis the proportion of eligible students enrolled in secondary and primary schools in 1960 (positively related to growth), per capita GDP at the beginning of the study's timeframe of analysis (negatively related to growth), and the price level of investment goods (also negatively related to growth).

According to the figure below, which shows a scatter plot of the relationship between the countries' TRUST and economic growth rates, the relationship appears to be positive (i.e., if you were to draw a line through the scattered points that represents a likely trend, the trend line would have a positive slope).



([Knack and Keefer 1997](#))

The table below presents the authors' empirical results based upon different specifications for ordinary least squares (OLS) regression equations:

Equation	1	2	3	4
Method	OLS	OLS	OLS	OLS
Dependent variable	Growth 1980–1992			
Constant	−0.935	−10.476	−9.593	−2.829
	(1.280)	(4.730)	(4.520)	(1.895)
GDP80	−0.361	−0.273	−0.375	0.152
	(0.131)	(0.126)	(0.127)	(0.274)
PRIM60	6.192	5.930	7.061	4.818
	(1.051)	(1.164)	(1.224)	(1.709)
SEC60	2.194	3.457	1.648	1.256
	(1.632)	(1.543)	(1.485)	(1.930)
PI80	−3.693	−3.117	−3.535	−3.930
	(0.867)	(1.100)	(0.935)	(0.755)
TRUST	0.082		0.076	0.192
	(0.030)		(0.030)	(0.060)
CIVIC		0.272	0.207	
		(0.098)	(0.092)	
TRUST*GDP80				−0.013
				(0.006)
Adj. R^2	.55	.44	.56	.60
SEE	1.37	1.52	1.35	1.29
Mean, D.V.	1.45	1.45	1.45	1.45

(Knack and Keefer 1997)

The social capital variables exhibit a strong and significant relationship to growth. For example, in Equation 1, the estimated coefficient for TRUST is positive (0.082) and statistically significant (due to its relatively low standard error of 0.030 in parenthesis). As Knack and Keefer explain, *TRUST*'s coefficient indicates that a ten-percentage-point increase in *TRUST*'s score is associated with an increase in economic growth of four-fifths of a percentage point. Similarly, according to *CIVIC*'s estimated coefficient, each four-point rise in the 50-point *CIVIC* scale in Equation 2 is associated with an increase in economic growth of more than one percentage point. When both social capital variables are entered together in Equation 3, their coefficient estimates drop slightly but remain statistically significant. Finally, the negative (and statistically significant) coefficient value on the interaction term *TRUST*GDP80* indicates that the effect of *TRUST* on economic growth is lower for countries with higher initial per-capita GDP levels at the beginning of the timeframe of analysis, in 1980 (represented by variable *GDP80*).

Therefore, it seems that Knack and Keefer's evidence of the extent to which trust and civic norms affect the welfare of a country supports the hypothesis that trust is indeed a form of social capital.

REPUTATIONAL EFFECTS

In Chapter 8, we learned of Fehr and Gächter's (2000) finding that Reputational Effects among a group of repeatedly partnered players in a laboratory-conducted, finitely-repeated Public Good Game are capable of mitigating free-riding behavior among the players (i.e., contribution levels that are repeatedly too low to adequately fund the public good). Concern about one's reputation among other players (for either strategic or non-strategic reasons) is a strong-enough incentive for players to voluntarily contribute at higher levels.

Curious about whether a Reputational Effect (or "indirect reciprocity") is capable of promoting large-scale cooperation in real world settings, Yoeli et al. (2013) designed a field experiment involving over 2,400 customers of a California utility company, Pacific Gas and Electric Company, in order to study the customers' levels of participation in a "demand-response program," called SmartAC, designed to prevent electricity blackouts (before getting into the proverbial weeds of the experiment, convince yourself that participation in a prevention program like this indeed fits the definition of a public good).[20] The authors' hypothesis is that the effects of indirect reciprocity are strong in a setting such as this.

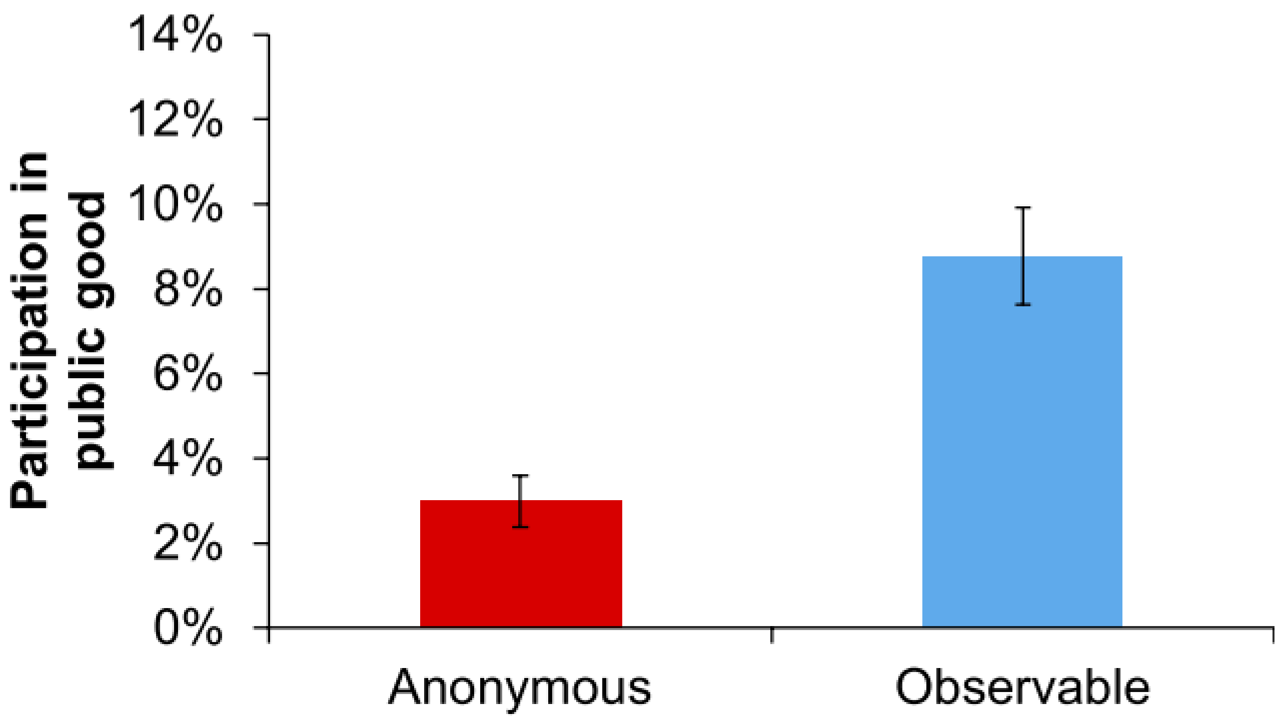
According to Yoeli et al., indirect reciprocity is based on repeated encounters in a group of individuals where my behavior toward you also depends on what you have done to others. We *Homo sapiens* have a relatively sophisticated social intelligence—we take a keen interest in who does what to whom and why. To be blunt, we gossip. And we are attuned to others' gossip about us. Indirect reciprocity enables us to track the good and bad behavior of others and, when it comes to contributing toward a public good, to use this information to incentivize cooperation.

The authors informed customers about the program via mailers. Sign-up sheets were simultaneously posted in a communal area near their home, usually by a shared mailbox kiosk. Those who signed up to participate in the program allowed the utility to install a device that remotely curbed their central air conditioners when necessary—on days with unusually high demand or in the case of an unexpected plant or transmission failure. In their primary manipulation, Yoeli et al. varied whether residents' neighbors could tell who had signed up for the program. They did so by dividing the publicly posted sheets between those requiring residents to print their name and unit number (observability treatment) and those providing a printed code number that did not reveal their identity (anonymous treatment). Note that participants in the observability treatment are susceptible to the effect of indirect reciprocity.

The figure below presents the experiment's general result. We see that observability tripled participation in the program, suggesting that reputational effects are indeed present in this public good experiment. Note that because the "whiskers" (|) at the top of the two boxes do not overlap with each other, the difference between the participation rates is statistically significant.[21]

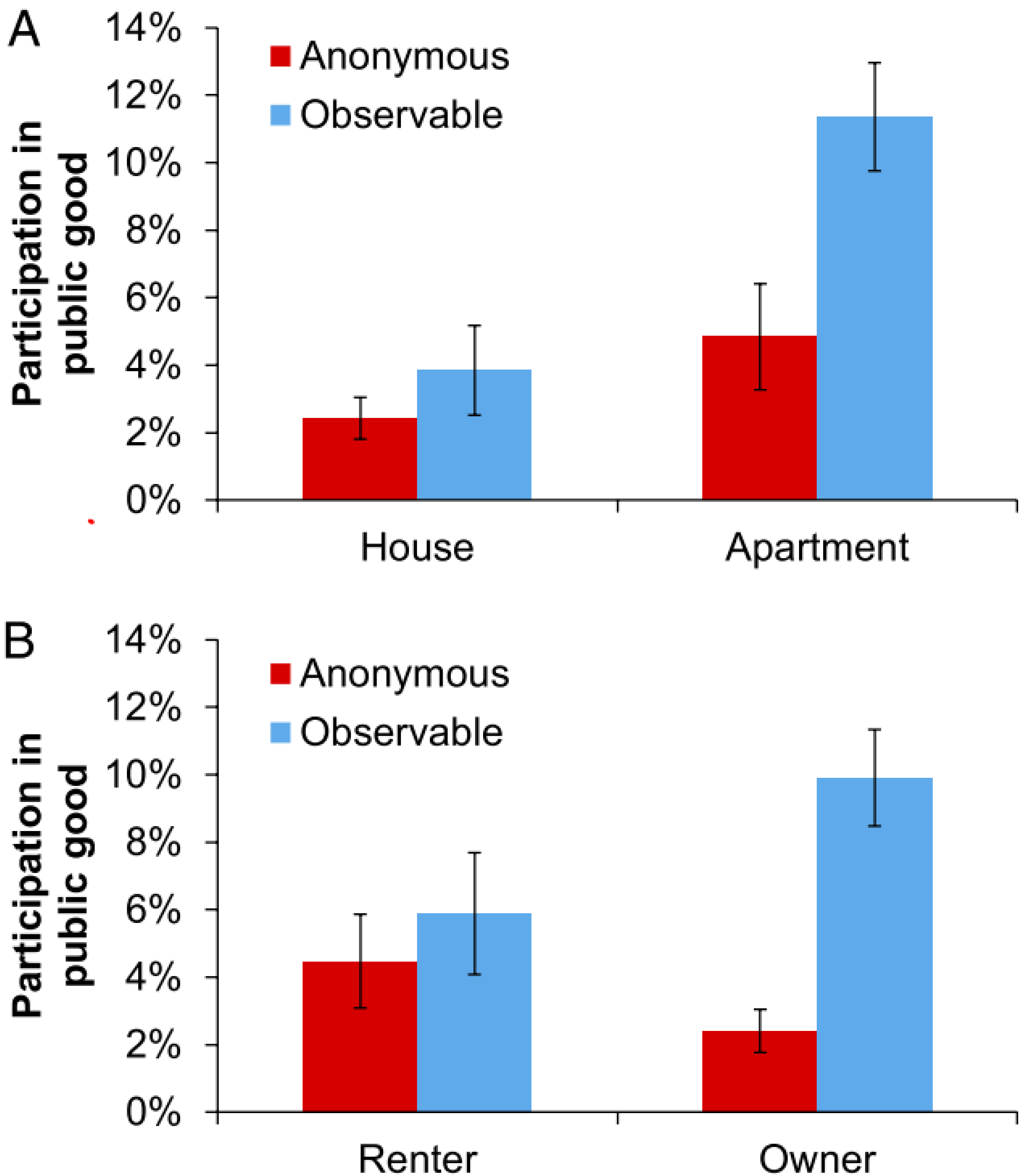
20. To help convince you, note that demand response programs are voluntary programs in which people allow their utility company to remotely restrict their electricity consumption during peak hours and thus reduce the risk of a blackout across the service area. To do so, the utility usually installs a remote switch in-line with the circuitry of an appliance such as a water heater or air conditioner. Excessive electricity usage during peak hours reduces grid reliability, drives up energy costs, increases the risk of blackouts, and harms the environment.

21. According to Yoeli et al., the effect of the observability treatment (measured in dollars) was over seven times that of a $25 incentive payment, which is what the utility had been offering before the experiment. The authors estimate that the utility would have had to offer an incentive of $174 to match the participation rate achieved via their observability treatment.



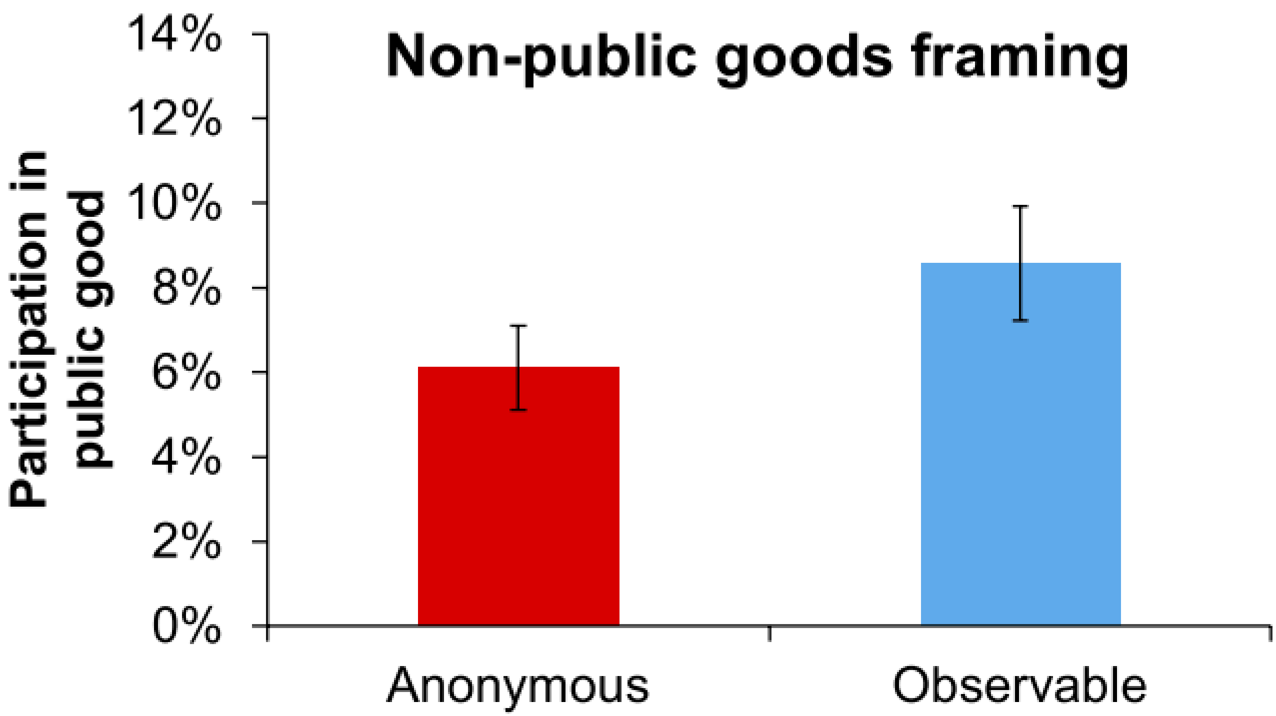
(Yoeli et al. 2013)

Charts A and B below dissect these results a bit further. In Chart A, we see that the observability treatment increased participation more in apartment buildings where residents are more likely to interact with their neighbors in public spaces, and sign-up sheets were posted in especially conspicuous locations, as compared with row houses or individual homes where neighbors are less likely to interact and sign-up sheets were less easily visible (note the lack of statistical significance for those living in homes—the whiskers overlap with each other). In Chart B, we see that the observability treatment increased participation more among those who own their homes/apartments relative to those who rent (note the lack of statistical significance for renters). The authors suggest that renters are more transient and therefore less likely to invest in long-term relationships with their neighbors.



(Yoeli et al. 2013)

On a final note, Yoeli et al. provide evidence that indirect reciprocity among *Homo sapiens* is unique to public goods. Their hypothesis is that choosing not to participate in a demand response program should carry the threat of social sanctions only if participation is considered to be for the public good. To test their hypothesis, the authors solicited an additional 1,000 customers with exactly the same treatments as described above, except that the informational materials the customers received ahead of time to entice them to participate in the demand response program were stripped of any language that framed blackout prevention as a public good. In the figure below, we see that, relative to the first figure above, the effect of indirect reciprocity is dramatically reduced among participants who did not receive the public good framing.



(Yoeli et al. 2013)

In the end, Yoeli et al.'s results suggest that *Homo Sapiens* are substantially more cooperative when their decisions are observable and when others can respond accordingly. The authors surmise that participants in their field experiment exhibited an understanding that having a good reputation is valuable in a public good setting and thus were willing to pay the cost of cooperation.

EMPLOYER-PROVIDED RETIREMENT SAVINGS PLANS

According to the Employee Benefits Research Institute (EBRI), less than a third of American workers feel very confident that they have saved enough money to live comfortably in retirement, and 60% report that preparing for retirement makes them feel stressed. Among those workers participating in an employer-sponsored, defined-contribution retirement plan, 80% report feeling satisfied with their plan and two-thirds are confident in their ability to choose the best-available retirement investments for their perceived situations. Only one-third were auto-enrolled into their plan. Overall, more than 30% of retirees feel that they do not have enough money saved to last their entire lifetimes (EBRI, 2020).

It is commonly believed that workers who fail to join an employer-sponsored plan, or who participate in the plan at low levels, appear to be saving less than they should for retirement—a mistake *Homo economicus* would naturally avoid making. In explaining this suboptimal behavior among *Homo sapiens*, behavioral economists stress lack of self-control, which leads to time-inconsistent investment choices being made over the course of a worker's career due to procrastination or Status Quo Bias. One potential solution to this problem has been for employers to automatically enroll their employees into a default plan, which then requires the employee to "opt-out" if they wish to make any changes to the default savings portfolio at any time during their employment.[22] The question of how workers should adjust their savings rates and portfolio allocations over time to ensure they are saving appropriately to meet their expected retirement needs looms large.

To overcome this potential time-inconsistency problem, Thaler and Benartzi (2004) proposed a new retirement savings program called Save More Tomorrow (SMarT). The program's commitment mechanism is straightforward. People commit now (when they begin the program) to increase their savings rate later (each time they get a pay raise). In other words, workers could continue to procrastinate about saving more for retirement over time and, in the end, still save more. Beautiful!

Thaler and Benartzi implemented the SMarT program as a natural experiment at an anonymous, mid-sized manufacturing company. The authors found that roughly 80% of those workers who were offered the plan joined, and 80% of those who joined it remained in the plan through a targeted fourth pay raise. The average saving rates for SMarT program participants increased from 3.5% to 13.6% over the course of a monitored 40 months. Employees who accepted an alternative saving recommendation increased their saving rate to a lesser extent, and those who declined both the SMarT and alternative savings plans saw no increase in their savings rate over the 40-month period.

Thaler and Benartzi find that more than half (162 out of 315) of the company's employees given the opportunity to participate in the SMarT program chose to do so. At the time of their first pay raise, the average savings rate for SMarT participants was equal to the average for those employees who made no effort to even contact the company's financial consultant, but less than the average savings rate for those who did contact the consultant and chose to adopt the consultant's recommended rate of slightly over 9%. However, by the second pay raise, SMarT participants were saving at a higher rate than any other employee group, and the differential in rates increased over the course of the subsequent two pay raises. It seems the SMarT program was successful in overcoming the employees' time-inconsistency problem with respect to biting the proverbial bullet and saving for retirement. SMarT was indeed a smart way to nudge workers into saving more for their retirements.

PUBLIC RETIREMENT SAVINGS PLANS

In contrast to private retirement savings plans like SMarT, Thaler and Sunstein (2009) describe Sweden's launch of an innovative public retirement savings program in 2000 aimed at overcoming potential time-inconsistent behavior among the country's workforce. All workers were instructed to choose between a default (opt-out) program designed by the national government or their own customized (opt-in) investment portfolio. By 2006, only 8% of new enrollees were customizing their own portfolios. This suggests that a sizable percentage of Swedish workers either recognized their penchant—as *Homo sapiens*—for making sub-optimal time-inconsistent decisions when it comes to saving for retirement, or they simply procrastinated their way into the default program.

On average, individuals who chose their own customized portfolio invested more in equities (particularly in Swedish equities) than those choosing the default program. The default portfolio was more diversified, more heavily invested in index funds, and carried a lower fee. Most importantly

22. To the extent that they are averse to opting-out of the default savings plan, and enrollment into the plan is automated, Homo sapiens succumb to what's known as Automation Bias.

from the investor's perspective, the default portfolio earned less-negative returns during the first three years and markedly higher positive returns over the subsequent three-year period.

Skål (as they are fond of saying in Sweden) to all the default Swedish savers! They responded well to the nudge of saving more for retirement.[23]

THE DEADWEIGHT LOSS OF GIFT-GIVING

Ho Ho Ho, or Ha Ha Ha? That's the question Waldfogel (1993) set out to answer about the time-honored tradition of gift-giving (e.g., during Christmas, Hanukkah, Valentine's Day, Mother's Day, weddings, births, etc.). Is the spirit of gift-giving (Ho Ho Ho) strong enough on its own merits to outweigh the potential deadweight loss imposed on *Homo sapiens* gift-givers and gift-recipients as a result of the gifts given (Ha Ha Ha)? As Waldfogel points out, an important feature of gift-giving is that consumption choices are made by someone other than the final consumer. As a result, gifts may be mismatched with the recipients' preferences. According to the rational model of choice behavior, the best a *Homo economicus* gift-giver can do with, say, $10, is to duplicate the choice that the recipient would have made. Because he implicitly solves the problem of maximizing the recipient's utility, a *Homo economicus* gift-giver gives cash if his perception of the recipient's utility from the cash gift, say $10, exceeds his perception of the recipient's utility from a non-cash gift costing $10.

While it is possible for a gift-giver to choose a non-cash gift that the recipient ultimately values above the price paid by the giver (e.g., when the recipient is not perfectly informed about a gift that she really enjoys), when it comes to *Homo sapiens* gift-givers, it is more likely the gift will leave the recipient worse off than if he had made his own consumption choice with an equal amount of cash. In short, gift-giving among *Homo sapiens* is a potential source of deadweight loss (terminology economists use to denote inefficiency) when the costs of something (in this case, gifts paid for by gift-givers) outweigh its associated benefits (recipients' valuations of their gifts plus the value gift-givers derive from the act of gift-giving itself).

Waldfogel estimates the deadweight loss of holiday gift-giving based upon surveys given to a group of Yale undergraduate students. He ultimately finds that holiday gift-giving results in deadweight loss ranging from 10% to a third of the value of gifts given. Non-cash gifts from friends and significant others are found to result in the least amount of deadweight loss, while those from members of the extended family result in the most. Given that holiday expenditures in the US in the 1990s averaged $40 billion per year, this would suggest a deadweight loss ranging from $4 billion to over $13 billion per year.[24]

Waldfogel's field experiment consisted of a series of two surveys administered to roughly 100 students over the course of three months. In the first survey (completed after the Christmas season

23. As you might imagine, nudging Homo sapiens via opt-out programs such as Sweden's public retirement savings program has received much attention (and been put into practice quite extensively) during the past few decades. Indeed, you will be learning about how default options have been used to save lives in the section below entitled Can Default Options Save Lives? As an example of how opt-out programs have been tested in randomized clinical trials, Montoy et al. (2016) conducted a natural experiment with patients in the emergency wing of an urban hospital. The authors found a statistically significant difference between patients agreeing to be tested for HIV under an opt-out agreement as opposed to an opt-in agreement (66% vs. 38% of patients agreeing to participate in the testing, respectively), a difference they call the Opt-Out Effect. Interestingly, the Opt-Out Effect was significantly smaller among those patients reporting high HIV risk behaviors.

24. Three years later, and using a different experimental design, Solnick and Hemenway (1996) reported a welfare gain (as opposed to deadweight loss) associated with gift-giving. The average subject in their experiment valued her gifts at 214% of the price paid by her gift-givers!

in January of 1993), the students were asked to estimate the total amounts paid by their respective gift-givers for all of the holiday gifts they received the previous month. Students were asked to place a value on each of their gifts based upon their hypothetical willingness to pay (WTP) for each gift and whether they later chose to exchange any of their gifts. The second survey (completed in March 1993) gathered additional data on each respondent's individual gifts listed in the first survey. The second survey asked respondents to describe each of their gifts, identify the givers' ages and relationships to the recipient (i.e., parent, aunt or uncle, sibling, grandparent, friend, or significant other), estimate the prices that the givers paid for the gifts, and indicate whether the gifts were ultimately exchanged. The gift description allowed the gifts to be divided into three categories: cash, gift certificates, and non-cash gifts. Perhaps most importantly, the students were again asked to place a value on each of their gifts, but this time based upon their hypothetical willingness to accept (WTA) payment for giving the gifts up.

In Survey 1, Waldfogel finds that students estimate that friends and family paid an average of roughly \$438 for the recipients' total gifts, but the students express an average WTP (or value) of only \$313 for the same gifts. The ratio of average WTP to average price paid (71.5%) suggests an average deadweight loss of roughly one-third of the value of all gifts given. Results from Survey 2—based upon the students' WTA values rather than WTP—suggest a deadweight loss closer to 10% of the value of all gifts given. Recall from Chapter 5, *Homo economicus and the Endowment Effect*, that we generally expect WTA values to exceed WTP values, which could explain Survey 2's lower estimates of deadweight loss from gift-giving.

Waldfogel goes on to report that aunt/uncle and grandparent gifts were the most likely to be exchanged, at rates of just under 21% and just over 13%, respectively. Ten percent of non-cash gifts received from parents were exchanged, as were roughly 7% of gifts from siblings and friends. A negligible number of gifts received from significant others were exchanged. Deadweight losses are larger for gifts given by extended family than by the immediate family, and losses increase with the age difference between the giver and recipient.

Recall that Waldfogel's deadweight-loss estimates were based upon hypothetical WTP and WTA values elicited from two survey instruments. List and Shogren (1998) put Waldfogel's findings to the test by instead eliciting valuations of an individual's gifts using an actual (i.e., real) "random nth price auction" in an effort to reduce potential Hypothetical Bias associated with Waldfogel's WTP and WTA estimates. As List and Shogren describe it, the auction works as follows:

1. For each gift received, an individual states his total value to sell the gift (i.e., states his WTA).
2. All gifts for a given individual, $g_i, i = 1, \ldots, I$, where I equals both the given individual's last gift and his total number of gifts, are then pooled together to create the set of total gifts across all individuals, $\sum_{i \in M} g_i$, where M represents the total number of individuals participating in the experiment.
3. The set of total gifts, G, is then rank-ordered from lowest to highest gift number across the M individuals.
4. The experimenter then selects a random number, n, uniformly distributed between 2 and 21 (2 was the lowest and 21 the highest number of gifts received by the individuals participating in the experiment).
5. The experimenter then purchases (with real money) the $(n - 1)$ lowest total value (i.e., lowest

WTA) gifts overall and pays the nth lowest total value for each gift. For example, suppose n = 6. Then, only the five lowest-valuation gifts overall (across the M individuals) would be purchased at the sixth lowest WTA value.

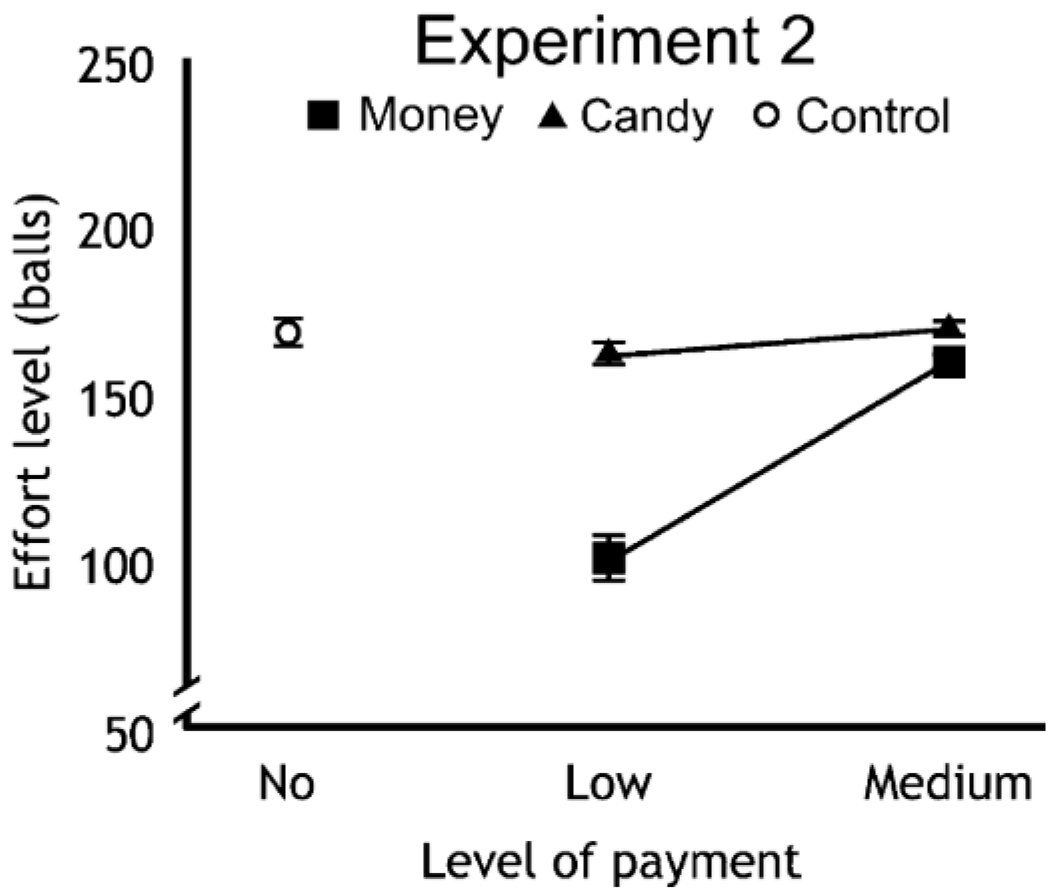
Complicated? A bit. But it seems a small price to pay (no pun intended) to mitigate potential hypothetical bias. List and Shogren go on to estimate a welfare gain associated with gift-giving—their average percentage yields range between 121% and 135% (as opposed to Waldfogel's 66% and 87% from the table above). Hence, it appears that evidence concerning gift-giving is context-specific—it depends upon how a given experiment is designed or framed. Hypothetical surveys suggest the existence of a deadweight loss. Real auctions suggest the existence of welfare gains. It seems we've been framed again by *Homo sapiens*.

THE BEHAVIORAL AND PSYCHOLOGICAL EFFECTS OF MONEY

As Heyman and Ariely (2004) point out, *Homo sapiens* often solicit help with tasks such as moving their possessions to a new residence, painting a room, or taking care of their kids. When we ask for help, we may wonder whom to approach and how best to motivate him or her. Should we ask a professional or a friend? If we ask a friend, should we offer compensation? If so, how much should we offer, and what form of compensation would be most effective? Would cash or personal gifts provide a stronger incentive? Using monetary payments causes participants to invoke monetary-market frames and norms. When money is not involved (i.e., there is either payment in the form of a gift or no payment is made at all), the market is perceived to be a social market invoking social norms. The authors discuss a set of experiments they designed to demonstrate that monetary vs. gift payments have material consequences for the payment-effort trade-off. Note that there is no such trade-off in the mind of *Homo economicus*. *Homo economicus* simply calculates the monetary value of the gift payment and thereby obviates any inherent distinction between the cash- and gift-payments.

In one experiment (Experiment 2), approximately 160 students each repeatedly dragged a computerized ball to a specified location on a computer screen. The software explained to the participants that a light gray circle (the "ball") would appear on the left-hand side of the screen and that their task was to drag as many of these balls as they could into a dark gray square on the right-hand side of the screen over the course of a three-minute period. Next, participants saw a screen that informed them of the payment they would receive (unless they had been randomly selected into the control condition of no payment). Those randomly assigned to the cash-payment treatment were paid in cash and those assigned to the gift-payment treatment were paid in an equivalent amount of Jelly Belly jellybeans.

Participants were not told the market price of the candy. The level of payment was either low (10 cents in the cash-payment treatment or five Jelly Bellies in the gift-payment) or medium ($4.00 in the cash-payment treatment or a half pound of Jelly Bellies in the gift-payment treatment). Results from this experiment are depicted in the figure below.



(Heyman and Ariely 2004)

We see four key results in this figure. First, the average participant's effort level (with respect to the ball-dragging task) in the cash-payment treatment increased significantly when the payment level increased from low to medium. Second, effort level in the gift-payment treatment is insensitive to the increase in payment level from low to medium. Third, effort level in the low-payment level of the cash-payment treatment is significantly below that of the no-payment control condition, but effort in the low-payment level of the gift-payment treatment is not. Lastly, the difference in the effort levels in response to the low level of payment in both the cash- and gift-payment treatments is statistically significant. In summary, these results support the distinction between monetary and social markets. In particular, they demonstrate that the decrease in performance from no-payment to low-payment conditions is found in monetary exchanges, but not in gift exchanges.

In another experiment, Heyman and Ariely tested the effects of monetizing the value of the gift payment (e.g., rather than valuing the low-payment gift as five Jelly Bellies, it was described as 10 cents worth of Jelly Bellies). The authors' prediction was that once the retail value of the candy was mentioned, the average participant's effort would be similar to that observed in the cash-payment treatment (i.e., the *Homo sapiens* participants would have no reason not to behave like *Homo economicus*). This is indeed what occurred, leading Ariely (2008) to state that "Once the bloom is off the rose—once a social norm is trumped by a market norm—it will rarely return" (page 85).

Ariely (2008) eloquently extrapolates the results of these experiments to a broader social context:

"If corporations started thinking in terms of social [markets], they would realize that these [markets] build loyalty and—more important—make people want to extend themselves to the degree that corporations need today: to be flexible, concerned, and willing to pitch in. That's what a social relationship delivers." (page 90)

Hence, in the less-predictable world of *Homo sapiens*, businesses must decide the extent to which they participate with their employees and customers in monetary and/or social markets.

As a follow-on to Heyman and Ariely's (2004) experiments exploring the payment-effort trade-off, Vohs et al. (2006) sought to understand the behavioral psychology underscoring the trade-off. In its most general terms, the authors' hypothesis is that money makes *Homo sapiens* feel self-sufficient and behave accordingly. When reminded of money, people desire to be free from dependency upon others and prefer that others not depend upon them. Vohs et al. designed several experiments to test this hypothesis from a variety of angles.

In one experiment, the authors found that participants (a sample of University of Minnesota students) who were reminded about money—both Monopoly money and real money—in the context of a series of word descrambling tasks worked longer at the tasks than participants in a non-money-primed control group before requesting help from the experimenter.[25] In subsequent experiments with different groups of students, Vohs et al. found that (1) participants in a high-money treatment worked significantly longer than participants in a low-money treatment before asking for help from another available participant, (2) participants in a money-primed treatment volunteered to help code fewer data sheets than did participants in the non-money-primed control condition, (3) participants in a high-money treatment volunteered to gather fewer pencils that had spilled onto the floor than did participants in a low-money treatment, and (4) participants in a money-primed treatment donated significantly less money to a university student fund than participants in the non-money primed control. Three final experiments tested the effects of money on social intimacy, desire to engage in leisure activities alone, and preference to work alone. As expected, participants who were primed with money ahead of time were subsequently less socially intimate and exhibited a stronger preference for engaging in leisure activities and working alone.

So yes, Vohs et al.'s experiments suggest that money makes *Homo sapiens* feel self-sufficient and behave accordingly.

PRICE AND THE PLACEBO EFFECT

Is it possible that the magnitudes of placebo effects experienced by *Homo sapiens* (e.g., through medical therapies or medications) are somehow influenced by the prices we pay for them? To investigate this possibility, Waber et al. (2008) studied the effect of price on a group of *Homo sapiens*' analgesic responses to placebo pills. Over 80 healthy volunteers in Boston, MA were recruited via an online advertisement to participate in a field experiment where each participant was informed by a brochure about a purported new opioid analgesic recently approved by the Food and Drug Administration. The opioid was described as similar to codeine but with a faster onset time. In reality, and not disclosed to the participants, the pill was a placebo. After randomization, half of the participants were informed that the drug had a regular price of $2.50 per pill ("regular price"), and half of the participants that

25. The descrambling task consisted of 30 sets of five jumbled words. Participants created sensible phrases using four of the five words. In the control and play-money treatment, the phrases primed neutral concepts (e.g., "cold it desk outside is" became "it is cold outside"). In the real-money treatment, 15 of the phrases primed the concept of money (e.g., "high a salary desk paying" became "a high-paying salary"), whereas the remaining 15 were neutral phrases. Participants in the play-money treatment were primed with money by a stack of Monopoly money in their visual periphery while completing the neutral descrambling task.

the price had been discounted to $0.10 per pill with no reason mentioned for the price discount ("low price").

The experiment followed the established approach for studying electrical shocks which were administered to the wrist and calibrated to each participant's pain tolerance level. After calibration, participants received the test shocks, rating the pain on a computerized visual analog scale anchored by the labels "no pain at all" and "the worst pain imaginable." Participants received shocks in 2.5-volt increments between 0 volts and their calibrated tolerances. Shocks at each intensity level were carried out twice for each participant (before and after taking the pill), and the change in reaction to the shock was assessed.

The authors found that, when informed of the regular price, slightly over 85% of the participants experienced pain reduction after taking the pill. This was a significantly higher percentage than the slightly over 60% of participants who reported pain reduction when informed of the low price. Waber et al. also found that for 26 of 29 intensities (from 10 to 80 V), average pain reduction was assessed as being greater for the regular-priced than the low-priced pill. Those informed of the regular price reported experiencing greater pain reduction beginning at roughly 25 volts (the authors report that the mean differences are statistically different for the shock intensities of 27.5 volts through 30 volts, 35 volts through 75 volts, and at 80 volts). In other words, Waber et al. found an abundance of evidence suggesting that *Homo sapiens* do indeed correlate perceived reductions in pain (as induced by placebo effects) with the placebo's per-unit price. Placebo effects are perceived to be more effective as they become more expensive. Ouch.

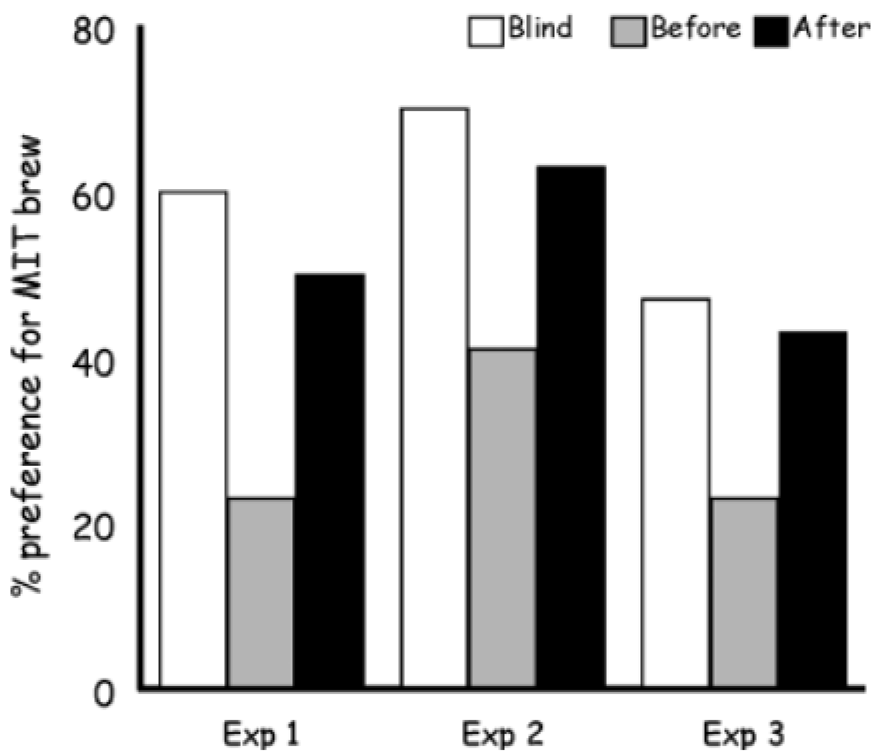
THE EFFECTS OF CONCEPTUAL INFORMATION ON THE CONSUMPTION EXPERIENCE

To what extent does conceptual (e.g., imaginary) information about a good and a consumer's expectations about the quality of that good influence the consumer's subjective experience of consuming the good? As early experiments with consumers demonstrated, *Homo sapiens*' preferences can indeed be influenced by conceptual information. For example, McClure et al. (2004) found in their experiments that Coca-Cola was rated higher when consumed from a cup bearing the Coca-Cola brand logo rather than from an unmarked cup. Wansink et al. (2000) similarly found that describing the protein of nutrition bars as 'soy protein' caused them to be rated as more-grainy and less-flavorful than when the word "soy" was removed from the description. However, as Lee et al. (2006) point out, none of these early experiments measured the extent to which information disclosure affected the consumption experience itself (i.e., the perceived tactile quality of the good). The experiments instead merely measured the consumer's retrospective interpretation of the experience.

To better answer the question of how conceptual information affects the consumption experience, Lee et al. conducted a series of field experiments. In each experiment, participants consumed two beer samples: one unadulterated sample and one sample of "MIT brew" containing several drops of balsamic vinegar, a beer flavoring that most participants found conceptually offensive. Participants were randomly assigned to one of three treatments. In the "blind treatment," the participants tasted the two beers without any information provided about the contents, and then indicated their preferences. In the "before treatment" they were told which beer contained balsamic vinegar prior to tasting it, after which they indicated their preferences. In the "after" treatment the respondents tasted the beers, were then told which of the beers contained vinegar, and then indicated their preferences. Note that because the information about the MIT brew concerns something considered conceptually offensive, the information itself is, by default, conceptual.

The authors point out that if the balsamic vinegar's presence solely affects preferences, the timing of the information should not matter, and preferences for the MIT brew should be reduced equally in the before and after treatments relative to the blind treatment (i.e., blind > before $\approx$ after). In contrast, if the information influences the consumption experience itself, preference for the MIT brew should be markedly lower in the before treatment than in the after treatment (i.e., blind $\geq$ after > before).

The experiments were conducted at two local pubs: The Muddy Charles and The Thirsty Ear. A total of approximately 400 patrons of these two pubs tasted two 2-oz. samples of beer each. One sample was of unadulterated beer (Budweiser or Samuel Adams) and the other of MIT brew. Participants in Experiment 1 merely indicated which of the two samples they liked best. In Experiment 2, participants also received a full (10-oz.) serving of the sample they preferred. In Experiment 3, the blind treatment was the same as in Experiment 2, but in the before and after treatments, participants received a full (10-oz.) glass of regular beer, some balsamic vinegar, a dropper, and the "secret recipe" ("add three drops of balsamic vinegar per ounce and stir"). The figure below depicts Lee et al.'s results:

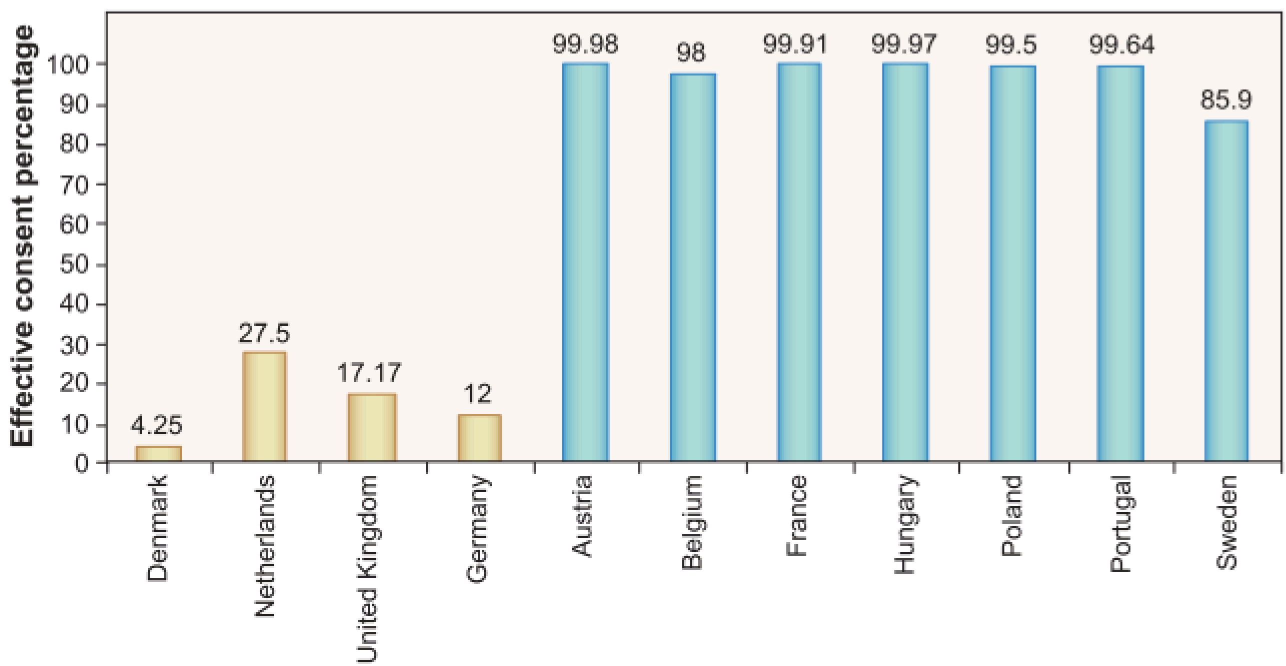


(Lee et al. 2006)

We see that in each experiment preference for MIT brew is (1) significantly higher in the blind treatment than in the before treatment, (2) significantly lower in the before treatment than in the after treatment, and (3) not significantly different across the blind and after treatments. In other words, blind $\approx$ after > before. Thus, the authors indeed find evidence that conceptual information—in this case, about something considered conceptually offensive—influences the consumption experience itself. Conceptual information can indeed alter *Homo sapiens*' expectations about the goods they consume.

CAN DEFAULT OPTIONS SAVE LIVES?

Johnson and Goldstein (2003) were motivated to ask this question because of glaring differences persisting between the US and several European Union nations when it comes to the role organ donations play in the saving of lives. In the US, thousands of patients die each year waiting for organ donations in spite of an oft-cited Gallup poll showing that (1) 85% of Americans approve of organ donation, (2) less than half of the American adult population have made a decision about donating, and (3) less than 30% have granted permission to harvest their organs by signing a donor card (Gallup, 1993). In the US, organ donation must be opted into via explicit consent, as it is in the United Kingdom, Germany, the Netherlands, and Denmark. To the contrary, in other European Union nations (e.g., Austria, Belgium, and France), organ donation must be opted out of. As Johnson and Goldstein show in the figure below, among European countries, the difference in effective consent percentages (ECPs) between explicit- and presumed-consent is stark:

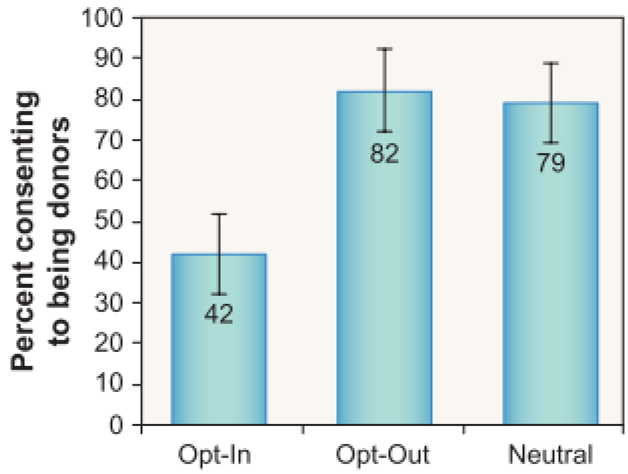


(Johnson and Goldstein 2003)

The ECP is the percentage of citizens who have opted in to donate their organs in explicit-consent countries, and the percentage who have not opted out in presumed-consent countries. In the figure, countries whose ECPs are represented by the gold bars are explicit-consent, and the countries whose ECPs are represented by the blue bars are presumed-consent. A picture is worth a thousand words here. On average, 60 percentage points separate the two groups.

To explain this difference, Johnson and Goldstein propose three possible reasons. First, citizens might believe that defaults are suggestions by their country's policymakers that imply a recommended action. In explicit-consent countries, the suggestion is to think hard about opting in, while in countries with presumed-consent the suggestion is to think hard about opting out. Second, since making a decision often entails effort and stress, whereas accepting the default is effortless, many people choose to avoid making an active decision about donating their organs. Third, default options often represent the status quo, and thus, change entails a trade-off. Due to loss aversion (which, as we know, is common among *Homo sapiens*), perceived losses associated with changing one's organ-donation status loom larger than equivalent gains.

The authors further investigate the effect of default options on donation rates by conducting an online experiment with over 160 respondents. The respondents were asked whether they would choose to become donors based upon one of three questions pertaining to different default options. In the question worded for the opt-in option, participants were told to assume that they had just moved to a new state where the default option was to not become an organ donor, and they were asked to confirm or change that status. The question for the opt-out option was worded identically, except that the default option was to become a donor. The third question was worded for a neutral condition, which simply required a respondent to choose whether to become a donor without any particular default option. Resulting ECPs are depicted in the figure below:



(*Johnson and Goldstein 2003*)

As the figure shows, the specific wording of the question had a dramatic impact. Stated ECPs were about twice as high when the respondent had to opt-out rather than opt-in. The ECP associated with the opt-out option did not differ significantly from the ECP for the neutral condition (without a specified default option provided). Only the ECP associated with the opt-in option, which represents the current practice in the US, was significantly lower than the ECP for the opt-out option.

The moral of this story, like that of the private and public retirement-savings stories encountered previously, is that merely framing a socially desirable choice as an opt-out decision can nudge *Homo sapiens* in the socially desirable direction.

REWARD VERSUS PUNISHMENT

Although not the domain of behavioral economists *per se*, the question of rewarding good behavior versus punishing bad behavior is a perennial one for anyone tasked with having to manage another's behavior or choices that determine a shared outcome—think parent-child, manager-worker, policymaker-citizen relationships. Do rewards for improved performance motivate better (i.e., nudge more) than punishments for mistakes?

Neuroscientists would argue that it depends. For example, Wachter et al. (2009) argue that rewards enhance learning in *Homo sapiens*, whereas punishment tends to improve motor performance. As Fryer et al. (2012) showed previously, rewards can work, particularly when framed as losses. Recall that the authors showed that teachers in Chicago (of K-8 students) who were paid in advance and asked to give the money back if their students did not improve sufficiently improved their students' math test scores. Teachers who were paid traditional subsidies for improved student performance did not improve their students' scores. And when it comes to reducing crime through greater punishment—in particular, higher arrest rates—Levitt (1998) showed that greater punishment can reduce certain types of crime, but not necessarily all types.

In one of the most highly cited field experiments involving the use of punishment, Gneezy and Rustichini (2000) found that punishment, if not administered at the correct level, can backfire, leading to more (not less) of the undesirable behavior. In their study of parents who were habitually late in picking up their children at Israeli daycare centers, a new fine levied on parent tardiness actually exacerbated the problem and ultimately led to adaptive behavior on the part of tardy parents. The authors concluded that penalties which are usually introduced into an incomplete social or private contract may change the information and perception among those being penalized regarding the environment in which they operate. The deterrence effect on behavior may therefore be opposite of what was expected.

Gneezy and Rustichini conducted their experiment at 10 daycare centers over a period of 20 weeks. In the first 4 weeks, they simply observed the number of parents who arrived late. At the beginning of the fifth week, they introduced a fine at six of the 10 daycare centers. The fine was imposed on treatment groups of parents who arrived more than 10 minutes late. No fine was introduced at the four other daycare centers, which served as the study's control groups. After the introduction of the fine, Gneezy and Rustichini observed a steady increase in the number of parents coming late. At the end of an adjustment period that lasted 2–3 weeks, the number of late-coming parents remained stable at a rate higher than during the no-fine period. The fine was removed (without explanation to the parents) at the beginning of the seventeenth week. In the following four weeks, the number of parents coming late remained at the same high level as the previous period, which was higher than during the initial four weeks. In other words, on average tardiness actually increased and was sustained among tardy parents with the onset of the fine, even after the fine was eventually eliminated.

One explanation for this perverse deterrence effect is simply that the fine was set too low. It could very well be that $3 per child was interpreted by some parents as signaling that tardiness was not considered by their daycare center to be a major problem. Paying what they consider to be a relatively low fine actually served to sanction their tardiness by relieving the guilt they otherwise might have felt in habitually arriving late to pick up their child. In this sense, the parents' willingness-to-pay (WTP) to relieve their guilt was greater than $3 per child. They were getting a deal!

Setting a fine or tax at the appropriate (or what economists call the "socially efficient") level is generally considered to be an antidote. For a recent example, Homonoff (2018) found that taxes (punishment) reduce demand for plastic grocery bags, whereas subsidies (reward) on reusable bags do not. Likewise, Haselhuhn (2012) found that a large fine boosts compliance more than a small fine, but the influence of paying both large and small fines decays sharply over time. This latter finding suggests that while *Homo sapiens* can react rationally to these types of nudges, their reference points die hard (or, should I say, evolve stubbornly over time). Because their effects can be transitory, penalties and rewards seen as being temporary are unlikely to establish a "new normal" that policymakers may be striving for.

Indeed, a growing body of research suggests that in certain circumstances, both penalties and rewards can backfire by crowding out *Homo sapiens'* intrinsic motivations and commitments to improve their behaviors, simply because the penalties and rewards are extrinsic (i.e., monetary) rather than intrinsic. For example, in their study of farmers in the La Sepultura Biosphere Reserve in Chiapas, Mexico, García-Amado et al. (2013) found that the more years a farmer has participated in a scheme where he is monetarily compensated for refraining from cutting down trees, hunting, poaching, or expanding the household's cattle herds, the more the farmer's stated preference for conserving the forest becomes financially driven. Further, a farmer's readiness to participate in future conservation efforts increasingly depends upon promised future payments. To the contrary, in other parts of Chiapas where the forest is communally managed, more time is initially required to galvanize farmer engagement, but their motivation remains centered on the intrinsic benefits of long-term forest conservation.

Beware the longer-term impacts of monetary incentives!

CONTINGENCY MANAGEMENT OF SUBSTANCE ABUSE

According to GBD 2016 DALYs and HALE Collaborators (2017), drug-use disorders are the 15th leading cause of disability-adjusted life years in high-income countries. Cocaine and amphetamines are the most commonly abused stimulants in people aged 15–64 years, with an annual prevalence of misuse among the global population of 0.38% and 1.20%, respectively (United Nations Office on Drugs and Crime, 2017). On the surface, these percentages may seem low, yet as the indirect effects on family members, friends, co-workers, and society at large of substance-abuse behavior are accounted for, virtually no one is unaffected.

As Degenhardt and Hall (2012) point out, patients addicted to stimulants experience a range of psychological and physical problems, including psychosis and other mental illnesses, neurological disorders, cognitive deficits, cardiovascular dysfunctions, sexually transmitted diseases, and blood-borne viral infections such as HIV and hepatitis B and C. Traditional approaches to recovery and rehabilitation, known as structured psychosocial interventions, tend to be expensive, embarrassing, difficult to access, and often ineffective (DynamiCare Health, 2020). These approaches eschew the use of explicit rewards and punishments, which, as we have previously learned, can be effective in altering a wide range of behaviors. So, why wouldn't reward schemes work against substance use disorders (SUDs) (e.g., in response to a patient maintaining drug-free urine samples over a specified period of time)?

As it turns out, clinical experiments with SUD reward schemes, commonly known as Contingency Management (CM), have a relatively long history, particularly in the short- and long-term treatment of people with cocaine and/or amphetamine addiction. Based upon their meta-analysis of 50 independent randomized control trials, De Crescenzo et al. (2018) conclude that CM, particularly in combination with community reinforcements (i.e., interventions involving functional analysis, coping-skills training, and social, familial, recreational, and vocational reinforcements), is the only intervention among traditional 12-step programs, Cognitive Behavioral Therapy (CBT), motivational interviewing, and non-contingent reward programs that increase the number of abstinent patients at the end of treatment (short-term), again at 12 weeks (medium term), and later still (longer-term).

Vincz (2020) reviews an ongoing telehealth recovery program undertaken by Horizon Blue Cross Blue Shield of New Jersey and DynamiCare Health involving approximately 300 patients struggling with SUD. Participants are required to pay a non-refundable $50 participation fee (which can be earned back within the first month through the program's reward system). They are then matched with a recovery coach (who are in recovery themselves) and receive breath and saliva testing equipment that works via a mobile app to support recovery remotely. The breath and saliva tests are conducted remotely through the app, relying on selfie video for verification. For staying sober and staying in treatment, members can earn monetary rewards worth up to $500 over the course of the 12-month program. The rewards come loaded on a smart-debit card which blocks access to bars, liquor stores, and cash withdrawals in order to protect the patient from risky spending.

The mobile app uses GPS technology to automatically check members into everything from medical appointments to Alcoholic Anonymous meetings, and record and reward patients for their participation in telehealth meetings and appointments. The app also contains a library of self-guided therapy modules based upon CBT. The short lessons teach crucial recovery skills such as how to deal with cravings, triggers, loneliness, and boredom.

If anything, this telehealth recovery program serves as an example of how a reward scheme paired with modern technology can be applied to one of society's most pernicious and persistent problems, and, to some extent, nudge us toward making healthier choices.[26]

F#!*ING PAIN MANAGEMENT

Cognitive scientists posit several reasons and motivations for why *Homo sapiens* swear. Swearing is an efficient way to convey emotion, it is cathartic, and it is an inexorable part of human evolution accompanying our innate fight-or-flight reactions (Bergen, 2016). But as Bergen and others point out, swearing can also serve as a mental analgesic, helping us cope with both physical pain and pain associated with social outcomes, such as ostracism.

Stephens and Robertson (2020) set out to test this assertion by generating two non-pre-existing "swear" words—"fouch" and "twizpipe"—that could conceivably be used in place of a conventional swear word—you guessed it, "fuck"—and to assess the pain-relieving effects associated with repeating these words in the context of an ice-water pain challenge.[27] A neutral word describing a standard wooden table (e.g., "solid") was used as a control condition to provide a reference against which to assess the effects of the conventional and new swear words. The authors hypothesized, *inter alia*, that the average *Homo sapiens*' pain threshold and tolerance levels would be higher for "fuck," "fouch," and "twizpipe" vs. the neutral word.

Approximately 100 students from Keele University participated in multiple trials of the experiment. For each student, the instructions for the ice water immersion were as follows:

> *In a moment, I would like you to fully immerse your nonpreferred hand into this ice water bath. While it is submerged, please repeat the word [INSERT AS APPROPRIATE] at normal speech volume and a steady pace, once every 3 seconds. While you have your hand in the water, I would like you to do TWO more things. First, please tell me when it becomes painful but don't take your hand out yet unless you have to. Second, please try and keep your hand in the water for longer, taking it out when the pain becomes unbearable.*

Timing began when the student's hand was fully immersed and stopped when her hand was fully removed from the water. Immediately after each submersion, participants immersed their hand in a room-temperature bath for three minutes prior to the next ice-bath submersion. Stephens and Robertson find that utterance of the swear word "fuck" not only induces significantly higher pain threshold and tolerance levels than the neutral word (measured by the number of seconds that the

26. See Hart (2013) for a discussion of innovative field experiments conducted at Columbia University with heroin and crack cocaine users that also test the efficacy of using monetary rewards to dissuade users from abusing these substances.

27. The size of the ice-water bath enabled a fully open hand to be immersed in water to a depth of approximately 120 mm.

average participant's hand is submersed in the ice bath), but also higher levels of pain threshold and tolerance than the made-up swear words "fouch" and "twizpipe". The authors find no statistical difference between the effects on pain threshold and tolerance of uttering "fouch" and "twizpipe" relative to the neutral word.

This suggests that when *Homo sapiens* decide to manage their pain with repeated utterances of a swear word, not just any word will do. Like the practiced eye of any connoisseur, the average *Homo sapiens*' ear can distinguish authentic from spurious swear words. It is unclear whether *Homo economicus*' ear is capable of such discernment.

WILLINGNESS TO ACCEPT PAIN (WTAP)

Yes, you've read that correctly—WTAP, or Willingness to Accept Pain. We're not talking about WTP (i.e., willingness to pay from Chapter 6 (recall *Homo economicus and the Endowment Effect*)). WTAP and WTP are two different things. For starters, while WTP is measured in dollars, WTAP is denominated in minutes (of pain tolerated). As such, WTAP is more similar to WTA (i.e., willingness to accept) than WTP. WTAP measures an individual's willingness to accept an additional dose of painful experience in exchange for a given monetary payment.[28] In Read and Loewenstein's (1999) field experiment, WTAP is defined specifically as the amount of time a subject is willing to keep her hand submerged in the ice water for $1, $3, and $5.

Read and Loewenstein subjected their experiment's participants (roughly 80 students and staff at the University of Illinois, Urbana-Champaign) to a 30-second ice-water pain challenge with the goal of measuring their WTAP with respect to their memories of the pain. Subjects either attended to the sensations of cold (sensation-focus condition, henceforth denoted as SENS) or were led to believe that the experiment was about manual dexterity (distraction, henceforth DIS). Subjects randomly assigned to the SENS condition were informed that the study was designed to assess the perception and memory of cold, while those assigned to the DIS condition were informed that the study was designed to assess manual dexterity under conditions of cold. In both conditions, subjects held a nut and bolt in their submerged hand and screwed and unscrewed the nut with their thumb and forefinger. Subjects in the DIS condition were told that their performance on this task was the focus of the study, while those in the SENS condition were not. WTAP was measured either immediately after pain induction (IMM) or following a delay of one week (DEL). Thus, there were four distinct experimental conditions: SENS/IMM, DIS/IMM, SENS/DEL, and DIS/DEL.

Read and Loewenstein's experiment spanned three consecutive weeks. In week 1, all subjects except those in a control group underwent pain induction. They grasped a large metal nut and bolt in their right hand and then immersed this hand into an insulated, two-liter bucket filled with ice water for 30 seconds. While their hand was immersed in the water, they undid the nut from the bolt and then tightened it back on using their thumb and forefinger, repeating the task until the experimenter instructed them to stop. Following pain induction, the delay groups were scheduled to return in a week and then dismissed.

At this point (in Week 1 for the subjects in the IMM and control conditions, but in Week 2 for the subjects in the DEL condition), all subjects stated their WTAP for the first time (WTAP1). They were

28. Recall the discussion of Less is More in Chapter 2. There we learned that an individual who conforms to the Peak-Ed Rule would prefer, say, three minutes of pain over two minutes of intense pain plus one minute of moderate pain over two minutes of intense pain (the same sequence minus the moderate pain) because the pain at the end of the longer sequence is lower than that at the end of the shorter sequence.

presented with the three money amounts ($1, $3, and $5) along with five time intervals for subsequent submersions of their hand in the cold bath (1, 3, 5, 7, and 9 minutes). Each of the 15 money-and-time combinations was written on a separate line, and subjects ticked off a box corresponding to "yes" (indicating that they were willing to immerse their hand in ice-cold water for that time in exchange for money) or "no" (indicating that they were not). Subjects were told that when they returned one week later, one of the money and time combinations would be chosen randomly and that their decision for that combination would "count." This meant that if they checked "yes" for a combination and that combination was randomly chosen in the draw, then they would be instructed to immerse their hand in the ice water for the specified period and would be paid the agreed upon amount for doing so. If they failed to hold their hand in the water long enough, they would not receive the extra money. If they had checked "no" on the chosen line, they would neither be asked to submerge their hand nor receive any extra payment. Although they were later given a chance to change their minds (WTAP2), at the moment when they made their first choices, subjects were led to believe that these choices would count.

The authors hypothesize that average WTAP1 estimates would be ordered in the following way (note that for WTAP1, smaller numbers mean that pain is judged to be greater):

$$WTAP1_{DIS/IMM} < WTAP1_{SENS/IMM} = WTAP1_{SENS/DEL} < WTAP1_{DIS/DEL}.$$

In other words, *Homo sapiens* judge pain assessed immediately after its occurrence to be greater than pain assessed after a delay of one week, all else equal. Among those assessing the pain immediately, those whose minds were distracted during the painful experience assess the pain to be greater than the pain suffered by those who were allowed to focus on the sensation of pain. Among those assessing the pain with a time delay, this relative assessment of the pain was reversed.

Read and Loewenstein find that those assessing the pain immediately and whose minds were distracted during the painful experience assess the pain to be greatest, while those whose minds were distracted but who assessed the pain after a time delay register the least pain. The authors concur that these results are statistically significant.

The Roman philosopher Seneca is credited with the aphorism, time heals what reason cannot. When it comes to the experience of physical pain, Seneca's aphorism seems to apply, particularly when *Homo sapiens* are able to distract their minds from the pain when it occurs. By contrast, *Homo economicus* would need no time delay to reason with their pain.

REDUCING URBAN HOMELESSNESS

As part of a grassroots campaign to fight homelessness, the city of Denver, CO installed "donation parking meters" where citizens can deposit loose change for community programs that provide meals, job training, substance abuse help, and affordable housing; change that would otherwise have been given to panhandlers (City of Denver, 2015). Approximately 100 of these meters were installed strategically on street corners where panhandling and pedestrian traffic occur at high levels. Each meter held up to $60 in change.

Denver's goal was to nudge residents and tourists to contribute $100,000 per year through the meters. The city also established a convenient way to text donations: text HOMELESSHELP to 41444. Charges appear on a donator's wireless phone bill. Jepsen (2019) reports that since Denver, CO and

Baltimore, MD pioneered their meters, approximately 50 US cities and two in Canada have installed donation meters. Most meters now accept credit card donations.[29]

The chief arguments in favor of the donation-meter approach to raising funds for worthy causes such as homelessness are (1) its convenience factor for both garnering donations and providing a depository for an individual's bothersome loose change, (2) the clever way in which it promotes awareness of homelessness and allows citizens to donate directly to the cause, increasing overall civic engagement, and (3) its potential deterrence effect on panhandling. It is well-known that convenience plays a key role in shaping the typical consumer's decision-making process (Kelley, 1958). Donation meters indulge the whims of modern-day *Homo sapiens* and can thereby provide a simple nudge where needed. The main argument against donation meters meant to reduce homelessness is that they discourage personal interactions that would otherwise be humanizing, inclusive, and promote greater mutual understanding.

REDUCING FOOD WASTE

Thaler and Sunstein (2009) report on a natural experiment conducted over the course of two days in 2008 by curious managers and students at Alfred University in New York City. The goal of the experiment was to test how much food waste could conceivably be saved if trays were removed from the university's cafeterias. The logic behind the experiment is simple. Since it is easy to load up a tray with extra plates of food that often go uneaten and extra napkins that go unused, eliminating the trays themselves, and thus forcing students to carry the plates in their hands to and from their tables, will help mitigate the waste *Homo sapiens* are prone to create in a market setting (where they face zero monetary expense for wasting food, the quantity of which is fully determined by their own choices).[30]

The managers and students found that, over the course of the two days, food and beverage waste dropped between 30% and 50%, amounting to 1,000 pounds of solid waste and 112 gallons of liquid waste saved on a weekly basis. Of course, the findings from the experiment were non-scientific, and therefore not generalizable to a wider population of cafeteria patrons.[31] Nevertheless, several other universities including New York University, the University of Minnesota, the University of Florida, Virginia Tech, and the University of North Carolina subsequently decided to designate some of their cafeterias tray-less.

It is interesting to note the difference between establishing tray-less cafeterias to reduce food waste on college campuses on the one hand, and re-purposing old parking meters to solicit donations to

29. In a nod to Denver's innovative spirit, the city of Steamboat Springs, CO recently installed re-purposed parking meters at local trailheads to encourage hikers to donate to trail maintenance on the spot (Associated Press, 2019).

30. We specify "fully" here because in an all-you-can-eat cafeteria, the students themselves choose how much food to bring to their table. In contrast, at sit-down restaurants, the students would only partially choose how much food is brought to the table. The owners of the restaurant determine the quantity of food on the plate that the waiter delivers, and then the student decides how much food to leave as waste at the end of the meal.

31. Had the Alfred University researchers wanted (and been able) to generalize their results, they would have needed to run the experiment over a longer period of time in order to account for seasonal and academic-scheduling effects (e.g., during exam weeks some students cope with the added stress of exams by adding plates of food to their trays as a way of comforting themselves). The researchers would also have needed to randomly assign some cafeterias to a treatment group (where the trays are removed) and a control group (where they are not) and periodically reassign the cafeterias from one group to the other throughout the semester. Further, they would need to periodically and randomly survey cafeteria patrons in both groups in order to identify those students who choose which cafeteria to dine at based at least partially upon whether that cafeteria is tray-less or not. This would allow the researchers to control for students who knowingly and purposefully avoid dining at the tray-less cafeterias to begin with, thus biasing the treatment effect downward.

reduce panhandling on the other (recall the section *Reducing Urban Homelessness*). In the former case, going tray-less serves as a punishment aimed at reducing a negative behavior many *Homo sapiens* have, unfortunately, habitualized by reducing the convenience factor associated with carrying plates of food on a tray. In the case of urban homelessness, installing donation meters is an attempt to increase a positive behavior that, unfortunately, not enough *Homo sapiens* seem to practice. This is accomplished by raising the convenience factor associated with donating what often seems to be troublesome amounts of spare change. The inconvenience of dealing with spare change is seemingly magnified in this age of ubiquitous credit card usage, not to mention the emergence of peer-to-peer payment apps such as Venmo, Skrill, and Zelle. Notwithstanding these different approaches used to reduce food waste and urban homelessness, it seems that simple societal nudges can be quite effective in helping to solve these types of problems.

REDUCING ENVIRONMENTAL THEFT

To test whether appealing to social norms can significantly reduce environmental theft from US national parks, Cialdini et al. (2006) conducted a field experiment where 2,700 visitors to Arizona's Petrified Forest National Park were exposed over a five-week period to signage admonishing against the theft of petrified wood. The signs conveyed information that appealed either to descriptive norms (i.e., the extent of other visitors' thefts) or injunctive norms (i.e., the levels of other visitors' disapproval of those thefts). The signs were combined with the park's existing signage which informs visitors that "Your heritage is being vandalized every day by theft losses of petrified wood of 14 tons a year, mostly a small piece at a time."

The descriptive-norm signage took one of two forms. One form (henceforth denoted D1) was negatively worded and accompanied by a photograph of three visitors taking wood from the park. The D1 sign read "Many past visitors have removed petrified wood from the park, changing the state of the petrified forest." The authors considered the combination of this signage and photograph to have a "strong focus" on the problem. The other sign (henceforth D2) was positively worded and accompanied by a photograph of three visitors admiring and photographing a piece of wood. The D2 sign read "The vast majority of past visitors have left the petrified wood in the park, preserving the natural state of the petrified forest." This signage-photo combination was considered to have a "weak focus" on the problem.

Similarly, the injunctive-norm took one of two forms. One form (henceforth l1) was supplicative and accompanied by a photograph of a visitor stealing a piece of wood, with a red circle-and-bar symbol superimposed over his hand. The l1 sign read, "Please don't remove petrified wood from the park," and the signage-photo combination had a strong focus. The other sign (henceforth I2) was also supplicative but was accompanied by a photograph of a visitor admiring and photographing a piece of wood. The I2 sign read "Please leave the petrified wood in the park," which, combined with the photograph, provided a weak focus. Hence, norms D1 and I1 provide a strong focus while norms D2 and I2 provide a weak focus.

The authors placed 300 marked pieces of petrified wood at each of the four different signage locations (D1, D2, I1, and I2) throughout the park. For their statistical analysis, they defined the key variable to be explained as,

$$\%theft = \frac{(\,pieces\ of\ marked\ wood\ stolen\ per\ signage\ location)}{300}.$$

Cialdini et al. found that injunctive norm I1 reduced theft the most, down to a theft rate of roughly 1.5%. In other words, a message with a strong focus on the problem that expresses disapproval of theft from the perspective of other visitors ("Please don't …") was quite effective at mitigating theft. To the contrary, descriptive norm D1 reduced theft the least (down to a theft rate of roughly 8%), suggesting that a message with a strong focus commenting on other visitors' behaviors and associated outcomes (but not explicitly expressing disapproval) was least effective. Hence, in the case of protecting environmental artifacts, emotional appeals incorporating explicit disapprobation, as opposed to mere comments on behavior and associated outcomes, seem to dispel the urge to steal among potential *Homo sapiens*. Similar to what we have seen with reducing food waste and urban homelessness and increasing personal savings rates, a nudge (in this case a carefully worded one) can help reduce environmental theft.

REDUCING LITTER

A common finding in the literature concerned with littering behavior among *Homo sapiens* is that people are more likely to litter in an already littered setting than in a clean setting.[32] This could be due to imitating others' behavior or because people perceive that their litter will do less damage in an already littered environment—two hypotheses suggesting that a person's propensity to litter is based upon what was previously defined as a descriptive norm.

To test these hypotheses, Cialdini et al. (1990) devised a series of novel field experiments to assess *Homo sapiens'* penchant for littering in public places. In Study 1, subjects encountered a large handbill tucked under the driver's side windshield wiper of their car in a parking garage. Seconds before reaching their cars, subjects in the randomly assigned treatment group witnessed a "confederate" littering the garage with his handbill (high-norm salience), and subjects in the control group witnessed a confederate who just walked by and did not litter his handbill (low-norm salience). Half of the parking garages had been (randomly) heavily littered beforehand by the experimenters with an assortment of handbills, candy wrappers, cigarette butts, and paper cups. Half of the garages were cleaned of all litter.

Overall, the authors found that subjects littered more in an already littered garage than in a clean garage. Further, when subjects observed a confederate littering in the littered garage they littered more, but littered less when observing a confederate littering in a clean garage. Specifically, subjects littered more in an already littered garage versus a clean garage in cases of both high-norm salience (54% vs. 6%) and low-norm salience (32% vs. 14%). However, while subjects in an already-littered garage littered more in the case of high-norm salience versus low-norm salience (54% vs. 32%), they littered less in a clean garage (6% vs. 14%).

Cialdini et al. conclude that the likelihood of an individual littering into an environment bearing various pieces of perceptible, extant litter will be described by a checkmark-shaped function. Little littering should occur in a clean environment. Still less should occur with a sole piece of litter in an otherwise clean environment, but progressively greater littering should occur as litter accumulates and the descriptive norm for the situation changes from anti-litter (low-norm salience) to pro-litter (high-norm salience).

In a second study (Study 2), subjects were college dormitory residents who found a handbill in their mailboxes. The environment in front of the mailboxes had been arranged so that it contained (a) no litter, (b) one piece of highly conspicuous litter (a hollowed-out, end piece of watermelon rind), or (c) a large array of various types of litter, including the watermelon rind. Again, a larger percentage of

32. Gladwell (2002) describes a similar theory, known as the Broken Window Theory, which states that a building's broken window can play a role in the proliferation of neighborhood crime.

subjects littered in an already littered environment (nearly 30%) than in a clean environment (11%). Interestingly, subjects littered less in a barely littered environment than in a clean environment (4% versus 11%). These results lead the authors to conclude that anyone wishing to preserve the state of a specific environment should begin with a clean setting so as to delay, for the greatest time possible, the appearance of two pieces of litter there. Those two pieces of litter are likely to begin a slippery-slope effect that leads to a fully littered environment and a fully realized perception that 'everybody litters here.'

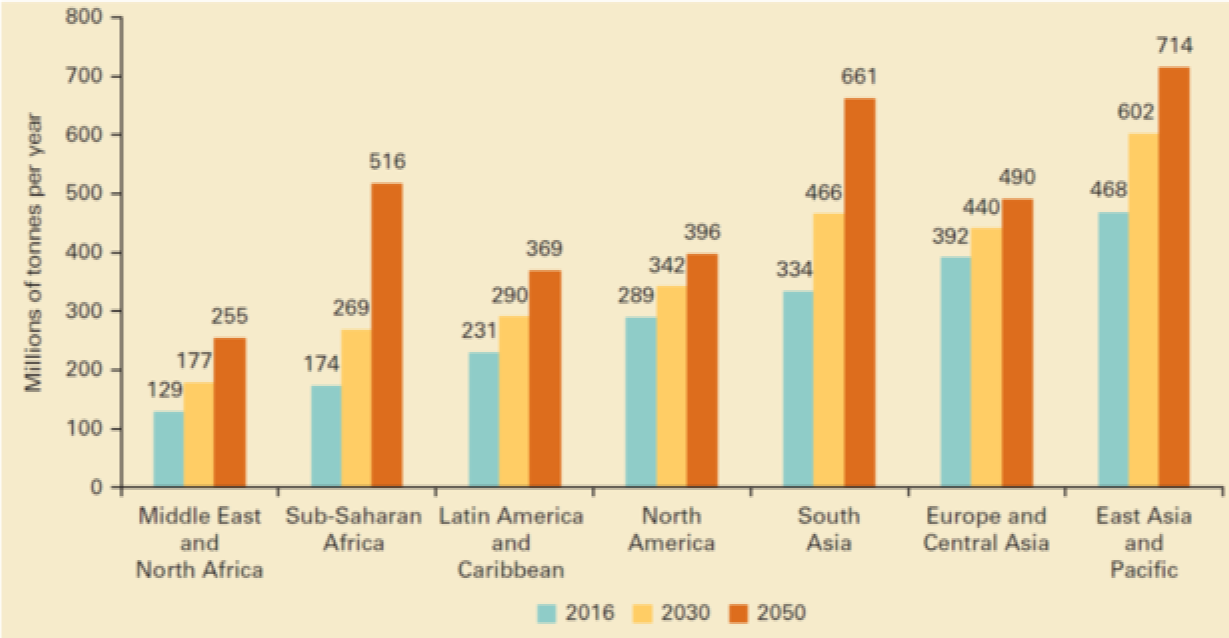
The results from Study 2 provoked Cialdini et al. to conduct a third study (Study 3) in order to test the strengths of the following injunctive norms (expressed in the form of a large handbill tucked under the driver's side windshield wiper of their car in a parking lot):

1. The handbill read, April is Keep Arizona Beautiful Month. Please Do Not Litter. (Anti-Littering Norm)
2. The handbill read, April is Preserve Arizona's Natural Resources Month. Please Recycle. (Recycling Norm)
3. The handbill read, April is Conserve Arizona's Energy Month. Please Turn Off Unnecessary Lights. (Turning Off Lights Norm)
4. The handbill read, April is Arizona's Voter Awareness Month. Please Remember That Your Vote Counts. (Voting Norm)
5. The handbill read, April is Arizona's Fine Art's Month. Please Visit Your Local Art Museum. (No Injunctive Norm)

As expected, the authors found that subjects (1) littered least after encountering the Anti-Littering Norm, (2) littered progressively more frequently as they encountered (equally normative) handbills 2-5, and (3) littered most when encountering no injunctive norm. In a proverbial nutshell, when it comes to reducing littering in public places, *Homo sapiens* generally respond as expected to descriptive norms, albeit in a non-linear (or check-marked) fashion. We respond linearly to increasingly targeted injunctive norms. In other words, as with environmental theft, *Homo sapiens* can be nudged away from littering with well-targeted appeals to social norms. In the case of littering, it helps to not let a location become littered in the first place.

GARBAGE IN, GARBAGE OUTED

Household garbage generation has accelerated quite considerably over the last few years in several regions of the world, inflicting substantial management costs and environmental burdens on citizens and their local governments. Higher wealth levels (resulting in higher consumption levels), higher urbanization rates, and more wasteful production methods are generally considered to be the driving forces behind this trend (Akbulut-Yuksel and Boulatoff, 2021; D'Amato et al., 2016). Sadly, as the bar chart below depicts, worldwide growth in municipal solid waste (MSW) is predicted to continue into the middle of this century, with particularly large increases occurring in Sub-Saharan Africa and South Asia.



(Kaza et al. 2018)

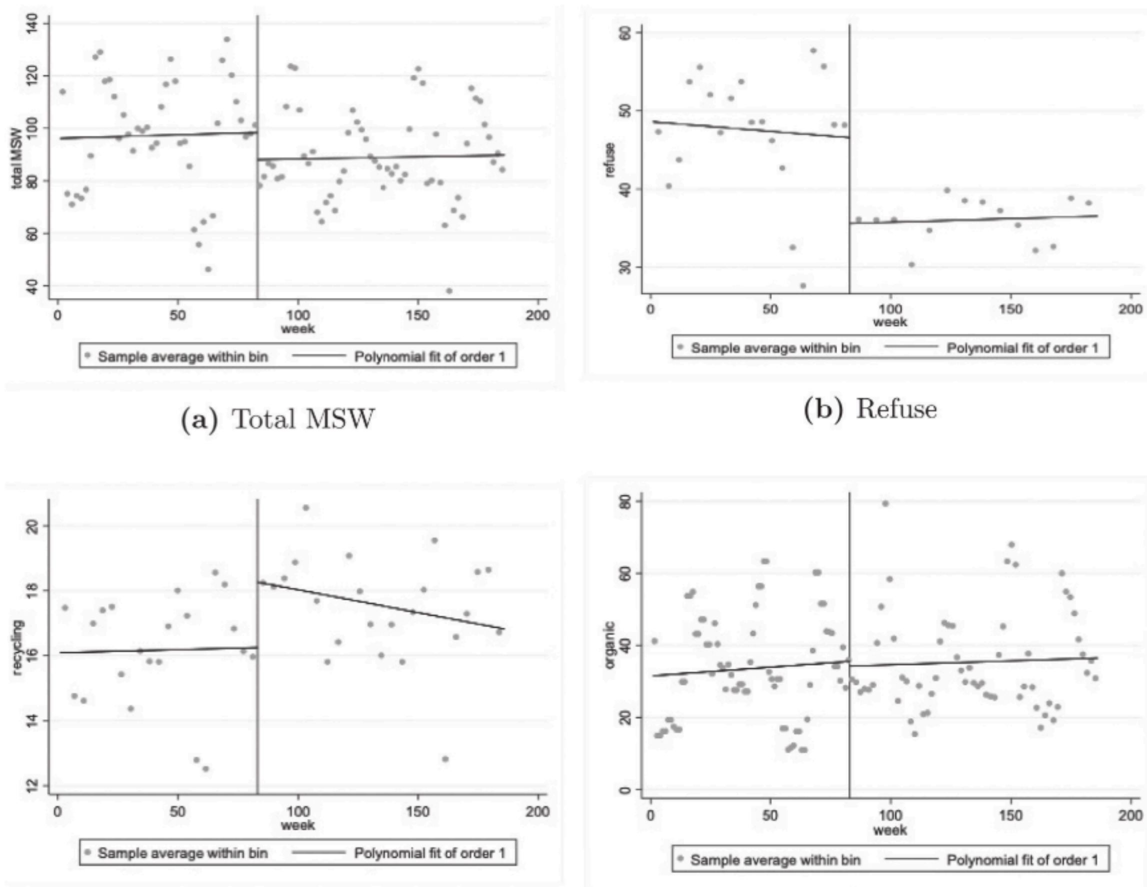
Canada is currently the world's largest producer of MSW per capita. At slightly more than 36 metric tons per person per year, Canadians generate roughly 10 tons more MSW per person annually than the next highest garbage producers, Bulgarians and Americans (Tiseo, 2021). Summiting a list like this is obviously not in any country's best interest—there are no kudos for reaching the top of the heap, so to speak. Is it therefore possible that those nations reaching the top will take the lead in reversing course?

Halifax is one Canadian city that apparently has. On August 1st, 2015, the city began providing a "green nudge" to citizens living in its urban core area with the introduction of the Clear Bag Policy, a policy designed to nudge households toward more responsible sorting of their waste, which, in turn, would result in an overall reduction in the total amount of waste generated. As Akbulut-Yuksel and Boulatoff point out, under the new policy, households were mandated to replace their black garbage bags, traditionally used for the disposal of their refuse, with clear, transparent bags. The Clear Bag Policy allowed households to put out the same number of garbage bags at the curb (six every other week), but all waste destined for the landfill was required to be disposed of in a clear bag (except for one dark bag permitted for privacy's sake). This allowed waste collectors to screen and refuse any bags containing materials that should otherwise have been diverted from the landfill, such as recyclables, food waste, and hazardous waste. Clear bags also made apparent to everyone, neighbors and passersby alike, a given household's waste-generation and disposal habits.[33]

To test the Clear Bag Policy's impact on a typical household's generation of MSW, Akbulut-Yuksel and Boulatoff designed a quasi-experiment spanning the period from January 6, 2014, to July 28, 2017, with January 6, 2014, to July 31, 2015, serving as the pre-treatment period and August 1, 2015, to July 28, 2017, serving as the post-treatment period. MSW data collected during this time span

33. As Akbulut-Yuksel and Boulatoff point out, Halifax households are required to sort waste in four ways: (1) recyclable containers (plastics, glass, and aluminum) are put in a transparent blue bag, (2) paper and cardboard are put in a separate bag, (3) organic food waste goes in a green bin provided by the city, and (4) the remaining waste (refuse) goes into garbage bags. Recyclable materials are collected each week, while garbage and organic waste are each collected every other week on opposite weeks (except in the summer months when, thank goodness, organic waste is collected on a weekly basis).

included the weight (in tons) of weekly recycling and bi-weekly garbage generated by households within the urban core area. The authors adopted a “regression discontinuity” design that exploits the differences in total waste, recycling, and refuse amounts in the weeks preceding and following August 1, 2015. Results are depicted in the figure below.



(a) Total MSW

(b) Refuse

(c) Recycling

(d) Organics

([Akbulut-Yuksel and Boulatoff 2021](#))

To begin, note that the vertical line in each panel (a)–(d) corresponds to the study’s 83rd week, the week of August 1, 2015 (when the Clear Bag Policy was implemented). In panels (a) and (b), we see statistically significant discrete drops at week 83 in total weekly MSW and landfilled refuse, respectively, generated by Halifax’s urban-core households—drops that are maintained for the remainder of the study period. In panel (c), we see a statistically significant increase in recycling; however, the increase occurring at week 83 is not maintained by the end of the study period. In panel (d), we see no statistical change in the amount of organic waste separated out for composting.

Akbulut-Yuksel and Boulatoff estimate that the Clear Bag Policy led to a 27% reduction in overall MSW, while increasing recycling by 15% compared to the pre-policy period. Their results also point to a short-term substitution effect between refuse and recycling (i.e., households became more responsible recyclers for a number of weeks after the policy was implemented). The authors found additional evidence suggesting that households located in neighborhoods with lower-than-average income and educational attainment exhibited larger improvements in their waste management and generation, thereby demonstrating that green nudges can affect household waste-management behavior differently across different socioeconomic groups. In the case of this particular study, the nudge exploited a household's innate concern about its reputation as a waste generator, not unlike the reputational effect we learned about earlier with respect to the SmartAC program designed to prevent electricity blackouts in Southern California.

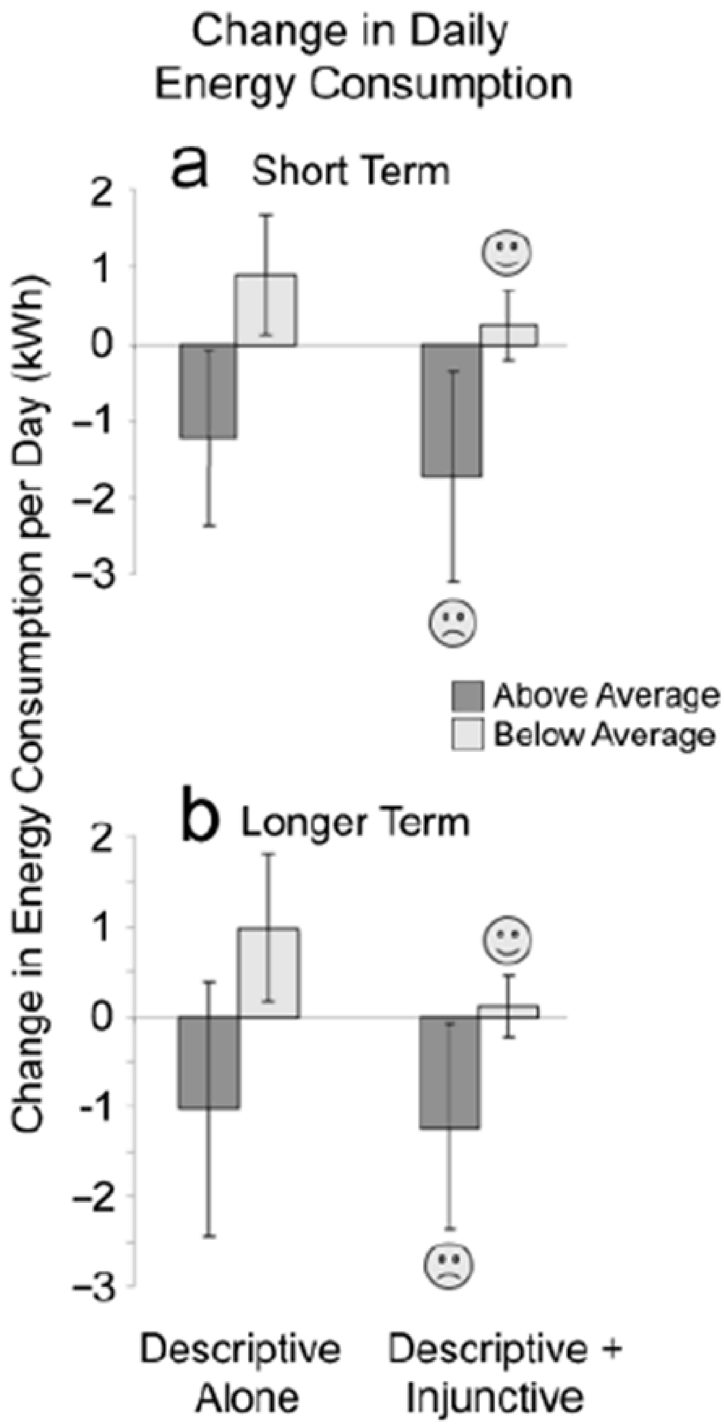
PROMOTING ENERGY CONSERVATION

Speaking of descriptive and injunctive norms, Schultz et al. (2007) conducted a field experiment with approximately 300 households located in a California community in which different messages were tested in promoting household energy conservation. Each message (included with a household's monthly energy bill) contained personalized feedback on the household's energy usage in previous weeks. For those households randomly chosen to receive a descriptive-norm message (henceforth denoted D) that included information about average household usage in the household's neighborhood for those previous weeks, the results were mixed. Households with higher-than-average usage reduced their energy usage while households with lower-than-average usage "boomeranged" by increasing their usage.[34] The authors claim that the former result indicates the constructive power of social norms, demonstrating that normative information can facilitate pro-environmental behavior. The latter result demonstrates the potentially destructive power of social norms, demonstrating that a well-intended application of normative information can actually serve to decrease pro-environmental behavior. Of course, *Homo economicus* households with lower-than-average usage would have responded to the D message by decreasing their energy usage, not boomeranging to higher usage.

Alas, when an injunctive-norm message (I) was added to message D (henceforth denoted D+I) and sent to a separate, randomly chosen group of households—where social approval (indicated by a smiley face emoji) or disapproval (indicated by a frowning face emoji) was provided based upon the household's usage relative to the neighborhood average—the boomerang effect disappeared. Hooray! Schultz et al. claim that this result demonstrates the potential reconstructive power of injunctive messages to eliminate the untoward effects of a descriptive norm. *Homo sapiens* households with lower-than-average usage just need a bit more nudging to reduce their energy usage.

The authors' specific results (in the form of box plots) are presented in the figure below:

34. Households were categorized as higher-than-average or lower-than-average based upon their energy usage during a two-week period prior to the commencement of the experiment.



(Schultz et al. 2007)

In this figure, Panel a presents results for the short term (where changes in usage are measured at the end of an initial one-week period) and Panel b for the long term (where changes in usage are measured at the end of a subsequent three-week period). The darker-shaded rectangles pertain to households with higher-than-average usage and the lighter-shaded rectangles pertain to households with lower-than-average usage. Further, the rectangles displayed on the left-hand sides of Panels a and b pertain to households that received message D, while the rectangles displayed on the right-hand side pertain to households that received message D+I. In cases where the segmented line (i.e., the "whisker") drawn through the middle of a rectangle does not extend beyond both the top and the bottom of the rectangle, the effect is considered statistically significant.

Hence, we see that, in the short term, households with higher-than-average usage who received the D message reduced their energy usage by a little over one kWh per day while households with lower-than-average usage increased theirs by a little under one kWh per day (this latter result demonstrates the boomerang effect). Also, in the short term, households with higher-than-average usage who received the D+I message again reduced their energy usage, this time by closer to two kWh per day. The D+I message provoked no statistically discernable effect on households with lower-than-average usage, thus eliminating the boomerang effect associated with the D message in the short term.

In the longer term, households with lower-than-average usage who received the D message continued using energy at an increased rate of roughly 1 kWh per day, but this boomerang effect was erased when the household received the D+I message. The short-term negative impact on households with higher-than-average usage who received the D message vanished over the longer term, but was sustained for households that had received the D+I message.

As with reducing environmental theft and littering, it seems that promoting energy conversation requires a well-targeted nudge.[35]

PROMOTING ENVIRONMENTAL CONSERVATION IN HOTEL ROOMS

Messages incorporating descriptive norms that promote prosocial behavior were put to the test in yet another context—as a means of reducing the use of fresh towels by hotel guests. To study the efficacy of messages including alternative descriptive norms (i.e., messages including information on other guests' behaviors), Goldstein et al. (2008) conducted two field experiments over separate 56- and 80-day spans with unwitting guests at a midsized, mid-priced hotel in the southwestern US. In the first experiment, data was collected on over 1,000 instances of potential towel reuse in 190 rooms. Two different messages urging guests' participation in the towel reuse program were printed on cards hanging from washroom towel racks, one of which each participating guest randomly received:

(1) A standard (control) environmental message focusing guests' attention on the general importance of environmental protection: "HELP SAVE THE ENVIRONMENT. You can show your respect for nature and help save the environment by reusing your towels during your stay," and

(2) A (treatment) descriptive norm message informed guests that a majority of other guests participate in the towel reuse program: "JOIN YOUR FELLOW GUESTS IN HELPING TO SAVE THE ENVIRONMENT. Almost 75% of guests who are asked to participate in our new resource savings program do help by using their towels more than once. You can join your fellow guests in this program to help save the environment by reusing your towels during your stay."

Below each of the respective messages on the cards were instructions on how to participate in the program: "If you choose to participate in the program, please drape used towels over the shower

35. Allcott (2011) studied the outcomes associated with sending descriptive-norm (D) messages to over half a million energy customers across the US. The author finds that households receiving a D message on average reduced their energy consumption by roughly 2 percent over a year. In a recently conducted natural experiment with 4,500 households in Southern California, Jessoe et al. (2020) find that semi-monthly D messages promoting water conservation (in the form of "Home Water Reports") spilled over into promoting short-lived reductions in electricity use during the summer months, when wholesale electricity prices and emissions are typically highest in the region.

curtain rod or the towel rack. If you choose not to participate in the program, please place the towels on the floor." Below the instructions, additional text informed the guests, "See the back of this card for more information on the impact of participating in this program." The information read, "DID YOU KNOW that if most of this hotel's guests participate in our resource savings program, it would save the environment 72,000 gallons of water and 39 barrels of oil, and would prevent nearly 480 gallons of detergent from being released into the environment this year alone?"

The authors found that, as predicted, the descriptive norm message yielded a significantly higher towel reuse rate (44%) than the standard environmental-protection message (35%).

In their second experiment, Goldstein et al. sought to investigate how hotel guests' conformity to such a descriptive norm varies as a function of the type of reference group attached to that norm (recall that the reference group referred to in the first experiment's descriptive-norm message was effectively the global norm of fellow hotel guests at large). The authors' hypothesis was that the closer individuals identify with their reference group and/or with their immediate surroundings, the more likely they are to adhere to a descriptive norm in making their own decisions.

For this experiment, the authors created five different towel hanger messages. The first two were the same standard and descriptive-norm messages used in the first experiment. The third was characteristic of a rationally meaningless and relatively non-diagnostic group—other hotel guests who had stayed in the guests' particular rooms. The last two signs conveyed norms of reference groups that are considered to be important and personally meaningful to people's social identities. Specifically, a fourth sign paired the descriptive norm with the reference group identity of "fellow citizens," and a fifth sign paired the norm with gender. Specifically,

> (1) The message for the same-room-identity descriptive norm message stated "JOIN YOUR FELLOW GUESTS IN HELPING TO SAVE THE ENVIRONMENT. In a study conducted in Fall 2003, 75% of the guests who stayed in this room (#xxx) participated in our new resource savings program by using their towels more than once. You can join your fellow guests in this program to help save the environment by reusing your towels during your stay."
>
> (2) The citizen-identity descriptive norm message stated "JOIN YOUR FELLOW CITIZENS IN HELPING TO SAVE THE ENVIRONMENT. In a study conducted in Fall 2003, 75% of the guests participated in our new resource savings program by using their towels more than once. You can join your fellow citizens in this program to help save the environment by reusing your towels during your stay."
>
> (3) The message for the gender-identity descriptive norm condition stated "JOIN THE MEN AND WOMEN WHO ARE HELPING TO SAVE THE ENVIRONMENT. In a study conducted in Fall 2003, 76% of the women and 74% of the men participated in our new resource savings program by using their towels more than once. You can join the other men and women in this program to help save the environment by reusing your towels during your stay."

The authors report that on average the four descriptive norm messages fared significantly better than the standard environmental message (44.5% vs. 37.2%). Thus, merely informing guests that other guests reused their towels induced participating guests to increase their towel reuse by more than if they had instead received a message focused explicitly on the general importance of environmental protection. Further, the same-room-identity descriptive norm message yielded a significantly higher towel reuse rate than the other three descriptive norm conditions combined (49.3% vs. 43%). Goldstein et al. conclude that towel reuse rates were actually highest for the participants' least-personally meaningful reference group (but most physically proximate). Therefore, when it comes to responding to descriptive norms about towel reuse, *Homo sapiens* tend to identify more with a spatially similar reference group than with a group sharing their personal characteristics.

FACE MASKS AND THE COVID-19 PANDEMIC

A question on the minds of Nakayachi et al. (2020) at the time of their study was why so many Japanese people decided to wear face masks during the pandemic, even though it was believed at that time that masks were unlikely to prevent them from getting infected with the virus? As the authors point out, wearing masks against COVID-19 was believed to be beneficial in suppressing the pandemic's spread, not through protecting the wearer from infection but rather by preventing the wearer from infecting others. Despite the belief that masks did not provide much protection, the custom of wearing masks prevailed in East Asia from the early stages of the pandemic, especially in Japan. Hence, Nakayachi et al. ask specifically, what are the psychological reasons prompting an individual to comply with a measure that is commonly believed not to provide any personal benefit? Sound familiar? Yes, we're talking about a public good here.

In their survey, the authors examined six possible psychological reasons for wearing masks, the first three of which involve individuals' perception of the severity of the disease and the efficacy of wearing face masks to reduce infection risks both for themselves and others. The first reason is an altruistic intention to avoid spreading the disease to others. Altruistic risk reduction to others is favorable for the whole of society. The second reason is self-interest in protecting oneself against the virus, even if wearing a face mask was at the time believed to be a misperception. If *Homo sapiens* are confident that masks will protect them against infection, they are likely to wear them. The third reason is perceived seriousness of the disease. The more an individual sees the disease as serious, the higher the person's motivation to take action.

The remaining three reasons involve other psychological driving forces. Reason four is that people may simply conform to others' behavior, perceiving a type of social norm in observing others wearing masks. Reason five is that wearing a face mask might relieve people's anxiety regardless of the mask's realistic capacity to prevent infection, and the sixth reason is that the pandemic has compelled people to cope as best they can. Wearing a face mask may be an accessible and convenient means to deal with the hardship.

Roughly 1,000 participants were recruited through electronic mail and accessed the designated website to participate in the survey. The survey was conducted between March 26 and 31, 2020. During this period, the total number of people infected with the virus in Japan increased from 1,253 to 1,887. Participants were asked about Covid-19 and the efficacy of masks, responding to six questions using a five-point Likert scale):

(1) Do you think your disease condition would be serious if you had COVID-19? (Severity)

(2) Do you think that wearing a mask will keep you from being infected? (Protection)

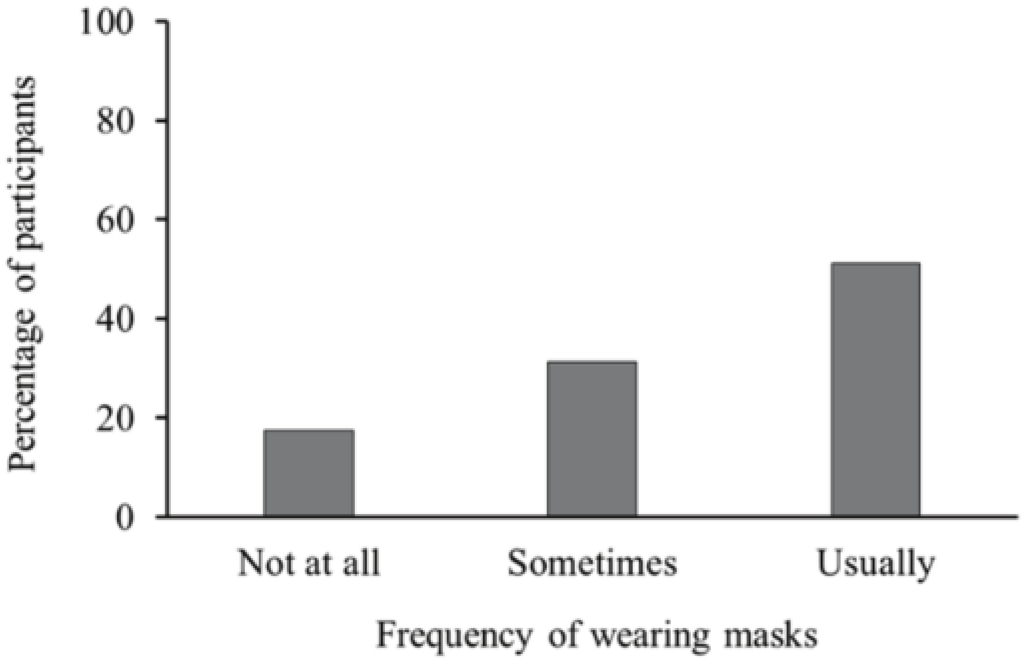
(3) Do you think that people who have Covid-19 can avoid infecting others by wearing masks? (Prevention)

(4) When you see other people wearing masks, do you think that you should wear a mask? (Social Norm)

(5) Do you think you can ease your anxiety by wearing a mask? (Relief)

(6) Do you think that you should "do whatever you can" to avoid COVID-19? (Impulsion)

Participants were also asked about their frequency of wearing masks during the pandemic, using a three-point scale. The figure below shows that more than half of the survey's participants usually wore masks from the beginning of the pandemic.



(Nakayachi et al. 2020)

The authors found a powerful correlation between perception of the social norm and mask usage—conformity to the mask norm was the most influential determinant. Feeling relief from anxiety by wearing a face mask also promoted mask use. By contrast, frequency of mask usage depended much less upon the participants' perceived severity of the disease and the efficacy of face masks in reducing infection risk, both for themselves and for others. These results lead Nakayachi et al. to conclude that effective nudging strategies against Covid-19 should appeal to social motivations among *Homo sapiens* such as the need to conform socially (at least as far as Japanese *Homo sapiens* are concerned). Further, the positive correlation between behavior and relieving anxiety by wearing face masks suggests that *Homo sapiens* consider subjective feelings rather than objective risks (i.e., when it comes to deciding the extent to which they will wear masks, *Homo sapiens* are prone to rely on an Affect Heuristic).

TEXT MESSAGING TO IMPROVE PUBLIC HEALTH

We have seen previously how providing nudges in the form of targeted messaging can help reduce environmental theft and litter and help promote energy conservation. Might similar forms of messaging be used to improve health outcomes among people whose treatments require frequent and consistent self-administration of drugs? In a fascinating field experiment, Pop-Eleches et al. (2011) test whether short message service (SMS) text reminders sent via cell phone to roughly 430 patients attending a rural clinic in Kenya are effective in inducing adherence to antiretroviral therapy (ART) for the treatment of HIV/AIDS.

Patients older than 18 years of age who had initiated ART less than three months prior to enrollment were eligible to participate in the study. Participants received Nokia mobile phones and were informed by the researchers that some participants would be randomly selected to receive daily or weekly text messages encouraging adherence to their ARTs. The participants were also informed that one of their medications would be dispensed in bottles with electronic caps enabling the researchers to monitor daily usage.

Participants were randomly assigned to one of four treatment groups or to a control group that received no text messages. One-third of the sample was allocated to the control group, and the remaining two-thirds of the sample were allocated evenly to each of the four treatment groups. As Pop-Eleches et al. explain, the four text-message treatments were chosen to address the different barriers faced by ART patients such as forgetfulness and lack of social support. Short messages (translated as, “This is your reminder”) served as simple reminders to take medications, whereas longer messages (translated as “This is your reminder. Be strong and courageous, we care about you.”) provided additional support. Daily messages were close to the frequency of prescribed medication usage, whereas weekly messages were meant to avoid the possibility that the more-frequent daily text messages would habituate the participants. Hence, the four treatments are as follows: short daily message, long daily message, short weekly message, and long weekly message. The messages were sent at 12 p.m. rather than twice daily (during actual dosing times) to avoid excessive reliance on the accuracy of the SMS software.

Participants were expected to return to the clinic once a month according to standard procedures. The electronic bottle caps were scanned monthly by the pharmacy staff. ART adherence was calculated as the number of actual bottle openings divided by the number of prescribed bottle openings for a given treatment period. The researchers’ primary determinant of patient adherence to the ART was whether the patient adhered at least 90% of the time during each of four 12-week periods of analysis. A secondary determinant of adherence was whether patients experienced a treatment interruption exceeding 48 hours during each period of analysis.

Pop-Eleches et al. find that the fraction of participants adhering to their ARTs at least 90% of the time in the two treatment groups receiving weekly reminders is significantly higher than the fraction of those adhering in the control group. Likewise, members of the weekly-reminder groups are significantly less likely than those in the control group to experience at least one treatment interruption during the entire 48-week follow-up period. Such is not the case for the members of the daily-reminder groups. Both the fraction of participants adhering at least 90% of the time and the fraction experiencing at least one treatment interruption are not significantly different for those receiving daily reminders than those in the control group. Lastly, compared with the control group, neither the long- nor short-message groups are better at adhering to their ARTs at least 90% of the time. However, the long-message group experiences marginally fewer treatment interruptions than the control group.

Again, we find that *Homo sapiens* can be fickle when it comes to the specific wording of messages meant to nudge them toward better personal outcomes. And in this case, we see that the frequency with which they are exposed to the messaging can influence the extent to which *Homo sapiens* are ‘nudgeable.’ Pop-Eleches et al. conclude that increased frequency of exposure to a message can lead to habituation, or the diminishing of a response to a frequently repeated stimulus. More frequent messaging might easily cross the line of intrusiveness and thereby be more likely to be ignored.[36], [37]

INVOKING FEAR AS AN AGENT OF CHANGE

It has long been believed that information alone seldom provides sufficient impetus for *Homo sapiens* to change both their attitudes and actions (c.f., Cohen, 1957). The information must not only instruct the audience but must create motivating forces which induce attitudinal and behavioral change. Leventhal et al. (1965) identified the arousal of fear as one potential motivating force for change and set out to test this hypothesis in the context of a field experiment that provided subjects with information encouraging inoculation against tetanus bacteria.

Spoiler alert: The arousal of fear resulted in more favorable attitudes toward inoculation and the expression of stronger intentions among the experiment's 60 subjects (who were seniors at Yale University) to get tetanus shots. However, actually taking action to get a shot occurred significantly more often among subjects who, in addition to having their fear aroused, also received information concerning a recommended plan of action to get the shot. Although actual decisions among subjects were unaffected by the fear factor in and of itself, some level of fear arousal was necessary for a subject to take action (i.e., to actually get inoculated). A recommended action plan was also not sufficient in and of itself for action to be taken by the subjects.

In Leventhal et al.'s study, fear-arousing and non-fear-arousing communications were used in recommending a clear action (getting a tetanus shot) which is 100% effective against contracting the disease. In addition, the perceived availability of a tetanus shot was experimentally manipulated by giving some subjects a specific plan to guide their action. It was hypothesized that subjects given a recommended action plan would choose to inoculate themselves at a higher rate. Most importantly, an interaction was anticipated between fear and action plan specificity: highly motivated

36. Wald et al. (2014) find that a combination of initial daily text messaging that slowly tapers off to weekly messaging is effective in improving adherence to cardiovascular disease preventative treatment among patients taking blood-pressure and/or lipid-lowering medications. In their field experiment, patients were randomly assigned to a treatment group that received text messages and a control group that did not. Texts were sent daily for the first two weeks, alternate days for the next two weeks, and weekly thereafter for six months overall. Patients in the treatment group were asked to respond (via reply text) on whether they had taken their medication, whether the text reminded them to do so if they had forgotten, and if they had not taken their medication. The authors found that in the control group 25% of the patients took less than 80% of the prescribed regimen compared to only 9% in the treatment group—a statistically significant improvement in adherence affecting 16 per 100 patients. Further, the texts reminded 65% of the treatment-group patients to take medication on at least one occasion and led 13% who had stopped taking medication because of concern over efficacy or side-effects to resume treatment.

37. Different countries' responses to the Covid-19 pandemic provide more recent evidence on the use of text messaging as a public health communication strategy. Considering a broad swath of countries' pandemic responses, Tworek et al. (2020) found that social messaging (of clearly stated, pro-social information aimed at strengthening democratic norms and processes) was an important component of each country studied. For example, in its social media campaign, Germany utilized Facebook and YouTube. The Federal Ministry of Health used Telegram and WhatsApp Covid-19 information channels as well as its own Instagram. New Zealand utilized the country's Civil Defense Alert System and the resources of the National Emergency Management Agency to communicate with citizens using mobile emergency alert messages, as well as Facebook Live video streams. Compared to its Nordic neighbors, the Norwegian government was most active on social media with both institutional and personal accounts, posting reflections and updates related to Covid-19 on Facebook, Instagram, and Twitter. Through emergency text messages and mobile applications (e.g., Corona Map), South Korean authorities managed to inform the public about the whereabouts of new patients. Social media (e.g., Facebook, Instagram, and KakaoTalk) were also widely utilized to disseminate vital public information and to build solidarity.

subjects—that is, those exposed to the fear-arousing messages—were expected to show the greatest attitudinal and behavioral compliance with the messages when a clear recommended plan of action was also provided to them.

As part of the experiment, subjects were randomly provided with one of four booklets (i.e., enrolled in one of four treatments), with each booklet containing two sections: a "fear section" dealing with the causes of tetanus and including a case history of a tetanus patient, and a "recommendation section" dealing with the importance of shots in preventing the disease. There were two treatments in each section: "high fear" and "low fear" in the fear section, and "specific recommendation" and "non-specific recommendation" in the recommendation section.

The high-fear treatments were distinguished by "frightening facts" about tetanus (as opposed to "non-frightening facts" in the low-fear form), "emotion-provoking adjectives" describing the causes and treatment of tetanus (as opposed to "emotion-non-provoking adjectives"), and graphic photographs of a specific case history (as opposed to non-graphic photographs of the case history). The specific and non-specific recommendation treatments included identical paragraphs on the importance of controlling tetanus by inoculation and illustrated by statistics that clearly demonstrated that shots are the only powerful and fully adequate protection against the disease. In addition, both recommendations stated that the university was making shots available free of charge to all interested students. The specific recommendation also included a detailed plan of the various steps needed to get a tetanus shot.

Two types of responses were measured for each subject. Immediately after reading their booklets, subjects completed a questionnaire regarding their attitudes, feelings, and reactions to the experimental setting, as well as any previous inoculations. In addition, a record was obtained of all subjects taking a tetanus inoculation. The records were checked by student health authorities, and a count was made of the subjects in each treatment who were inoculated. The dates for inoculation were also obtained.

As previously mentioned in the spoiler alert, the high-fear treatments were very successful in arousing fear and its attendant emotions. Subjects reported feeling significantly greater fright, tension, nervousness, anxiety, discomfort, anger, and nausea in the high- as opposed to the low-fear treatment. During the four-to-six-week period between the experimental sessions and the end of classes, nine of the 60 eligible subjects went for tetanus shots. Of the nine, four were in the high-fear, specific-recommendation treatment; four in the low-fear, specific recommendation; one in the low-fear, non-specific treatment; and none in the high-fear non-specific treatment. Thus, all else equal, subjects in the specific-recommendation treatments were more likely to get inoculated while subjects in the high-fear treatments were apparently not.

To test whether specific recommendations were sufficient in and of themselves (without either low- or high-fear stimuli) to impel the subjects to inoculate themselves, Leventhal et al. formed a control group consisting of 30 subjects who were exposed solely to the specific recommendation. The procedures for contacting and inculcating subjects were identical to those used in the original four treatments. Not one of the subjects availed himself of the opportunity to obtain an inoculation. Thus, the authors conclude that a specific recommendation alone is insufficient to influence actions or attitudes.

Such is the story for inoculating against a disease such as tetanus. Does this result concur with the previous results obtained for reducing litter, environmental theft, drunk driving, and increasing energy conservation? Such is *Homo sapiens'* varied responses to messaging.

INCOME TAX COMPLIANCE

In 1995 the Minnesota Department of Revenue (MDR) conducted a field experiment with 47,000 taxpayers (Coleman, 1996). The experiment tested alternative strategies to improve voluntary compliance with the state's income tax laws, including (1) increased auditing of tax returns with prior notice to taxpayers, (2) enhanced tax preparation services provided to taxpayers, (3) descriptive norm messages contained in letters sent to taxpayers, and (4) introduction of a more user-friendly tax form. The primary measures used to evaluate compliance were (1) a taxpayer's change in reported income between 1994 and 1995, and (2) a taxpayer's change in state taxes paid between 1994 and 1995.

MDR uncovered three sets of key results:

1. Lower- and middle-income taxpayers facing an audit reported more income and paid more taxes.[38] Increases were generally larger among taxpayers who had business income and paid estimated state taxes in 1993. Higher-income taxpayers had a mixed reaction to the threat of an audit—some responded positively, some negatively. The overall effect on their taxes was slight. Because they are expensive to conduct, audits are not particularly cost effective.

2. Enhanced tax-preparation services had no effect on reported income or taxes paid. Only 14% of taxpayers who were offered the expanded service availed themselves of it—slightly below the rate of taxpayers who had historically used traditional tax-preparation services at that time.

3. One of two messages contained in letters sent to taxpayers had a modest positive effect on reported income and taxes paid, which reinforces the argument that appeals to social norms increases responses. The letter read: "According to a recent public opinion survey, many Minnesotans believe other people routinely cheat on their taxes. This is not true, however. Audits by the Internal Revenue Service show that people who file tax returns report correctly and pay voluntarily 93% of the income taxes they owe. Most taxpayers file their returns accurately and on time. Although some taxpayers owe money because of minor errors, a small number of taxpayers who deliberately cheat owe the bulk of unpaid taxes (pages 5-6)."[39]

In a subsequent experiment, Alm et al. (2010) sought to uncover the extent to which uncertainty in how much tax is owed correlates with tax evasion. In the experiment, subjects accrued tax on income earned during a simple task. Subjects could claim both tax deductions and tax credits. If a subject decided not to file a tax return, she paid zero tax but missed out on claiming a tax credit. In the experiment's control group, subjects were made fully aware of the rules for claiming deductions and credits. In one treatment group, the uncertainty treatment, subjects had to guess the levels of deductions and credits they could claim. Only if they were audited would they learn how much tax they owed. In another treatment group, the information treatment, subjects had to guess how much

38. Lower- and middle-income taxpayers had a 1993 federal adjusted gross income below $100,000. High-income taxpayers had a 1993 federal adjusted gross income above $100,000.

39. Sadly, when it comes to nudging lower-income taxpayers to claim their Earned Income Tax Credits (EITCs) from the Internal Revenue Service (rather than pay taxes owed), Linos e al. (2022) find no response to a variety of different messaging approaches in their field experiments. This goes to show that nudges are certainly not failsafe.

income they should report to the tax authority, but they could press a button to learn exactly how much they should report. As expected, the authors found that both the filing and compliance rates were highest (although only slightly) among subjects in the information treatment. Similar to the provision of enhanced tax-preparation services in the MDR study, it appears that providing additional information on how much income to report on their tax returns does not, all else equal, compel the typical *Homo sapiens* taxpayer to comply with their taxpaying obligations.[40]

THE NOT-SO-GOOD SAMARITAN

It is helpful to know that the Judeo-Christian parable of the Good Samaritan has value in suggesting both personality and situational variables relevant to helping others. At least that is the conclusion reached by Darley and Batson (1973) in their innovative experiment with seminary students roughly 50 years ago. Using the Good Samaritan parable as their motivation, the authors presented the unwitting students with surprise, real-life encounters with a person in apparent distress and studied the students' responses. Surprisingly, a student's personality (or disposition) was unable to predict whether he would stop and offer assistance. In contrast, the extent to which a student was in a hurry as he came upon the person needing assistance (i.e., the situation) could explain the student's response—those in more of a hurry were less likely to stop and offer assistance.

The parable, which appears in the Gospel of Luke (Luke 10: 29-37 RSV), offers insight into the roles that both dispositional and situational effects are expected to play in summoning assistance from a passer-by:

> *"And who is my neighbor?" Jesus replied, "A man was going down from Jerusalem to Jericho, and he fell among robbers, who stripped him and beat him, and departed, leaving him half dead. Now by chance a priest was going down the road; and when he saw him he passed by on the other side. So likewise a Levite* [priest's assistant]*, when he came to the place and saw him, passed by on the other side. But a Samaritan, as he journeyed, came to where he was; and when he saw him, he had compassion, and went to him and bound his wounds, pouring on oil and wine; then he set him on his own beast and brought him to an inn, and took care of him. And the next day he took out two dennarii and gave them to the innkeeper, saying, "Take care of him; and whatever more you spend, I will repay you when I come back." Which of these three, do you think, proved neighbor to him who fell among the robbers? He said, "The one who showed mercy on him." And Jesus said to him, "Go and do likewise."*

As Darley and Batson point out, the Samaritan can be interpreted as responding spontaneously to the situation rather than being preoccupied with the abstract ethical or organizational do's and don'ts of religion as we might expect the priest and Levite to be. Hence, to the extent that the parable is relevant in the modern age, we should expect *Homo sapiens* to share more the disposition and situation of the Samaritan than that of the priest or priest's assistant to stop and offer assistance to someone in distress. Further, it is clear from the parable that the Samaritan had ample time on his hands to provide assistance. After binding the man's wounds, pouring on oil and wine, bringing him to an inn, and continuing to administer care there, the Samaritan still promised to return the next day to

40. Recent research by Heffetz et al. (2022) regarding compliance with parking tickets suggests that the success of simple nudges like these depends upon the recipient's characteristics. Reminder letters sent to parking-ticket recipients in New York City resulted in large differences in responses dependent upon the recipients' propensities to respond. In particular, low-propensity types (i.e., those facing significant late penalties or who come from already disadvantaged groups) reacted least to the letters.

check up on the man's recovery. This suggests that the Samaritan was not in a hurry at the time, and therefore his situation was even more amenable to stopping and offering assistance.

Darley and Batson coalesce these interpretations into three testable hypotheses, the first two of which correspond to situational effects and the third corresponding to dispositional effects:

1. *Homo sapiens* who encounter a situation possibly calling for a helping response while thinking religious and ethical thoughts will be no more likely to offer aid than persons thinking about something else.

2. *Homo sapiens* encountering a possible helping situation when they are in a hurry will be less likely to offer assistance than those not in a hurry.

3. *Homo sapiens* who are religious in a Samaritan-like fashion will offer assistance more frequently than those religious in a priest- or Levite-like fashion.

To test these hypotheses, the authors recruited 40 students at Princeton Theological Seminary to participate in a two-part field experiment. In the first part of the experiment, each subject was administered a personality questionnaire in order to identify the subject's respective "religiosity" type (e.g., whether a subject viewed religion as more a "means to an ends" in life, an "ends in itself," or as a "quest for meaning" in the subject's personal and social world—which is commonly believed to represent the Good Samaritan's religiosity). In the experiment's second part, the subject began experimental procedures in one building on campus and was then asked to report to another building for later procedures. While in transit, the subject passed a slumped "victim" planted in an alleyway. Unbeknownst to the student, measurements were taken of the degree to which he stopped and provided assistance to the victim.[41] Prior to being in transit, the student was told to hurry (at varying levels of admonition) to reach the other building. The student was also told the topic of a brief talk he was to give to a waiting audience after arriving at the other building. Some students were instructed to give a talk on the jobs in which seminary students would be most effective while others were instructed to give a talk on the parable of the Good Samaritan.

Darley and Batson found that subjects in more of a hurry (based upon the degree of admonishment provided by the experimenters) were (1) less likely to stop and offer assistance, but (2) once they stopped, less likely to offer less help than were subjects in less of a hurry. Whether the subject was going to give a speech on the parable of the Good Samaritan or job prospects for seminary students, it did not significantly affect his helping behavior in either of these two respects. These results confirm both Hypotheses 1 and 2. Regarding Hypothesis 3, the authors claim that religiosity played no role in either respect—either choosing to stop, or once stopped, choosing to provide a higher level of assistance.[42]

41. The victim was sitting slumped in a doorway, head down, eyes closed, not moving. As the student passed by, the victim coughed twice and groaned, keeping his head down. If the student stopped and asked if something was wrong or offered to help, the victim, startled and somewhat groggy, said, "Oh, thank you [cough]. . . . No, it's all right. [Pause] I've got this respiratory condition [cough]. . . . The doctor's given me these pills to take, and I just took one. . . . If I just sit and rest for a few minutes I'll be O.K. . . . Thanks very much for stopping though [smiles weakly]." If the student persisted, insisting on taking the victim inside the building, the victim allowed him to do so and thanked him.

42. Surprisingly, Table 2 in the article suggests that subjects who, all else equal, view religion as more a "means to an ends" in life were less likely to stop and assist the victim, and among those who did stop, they provided a lower level of assistance.

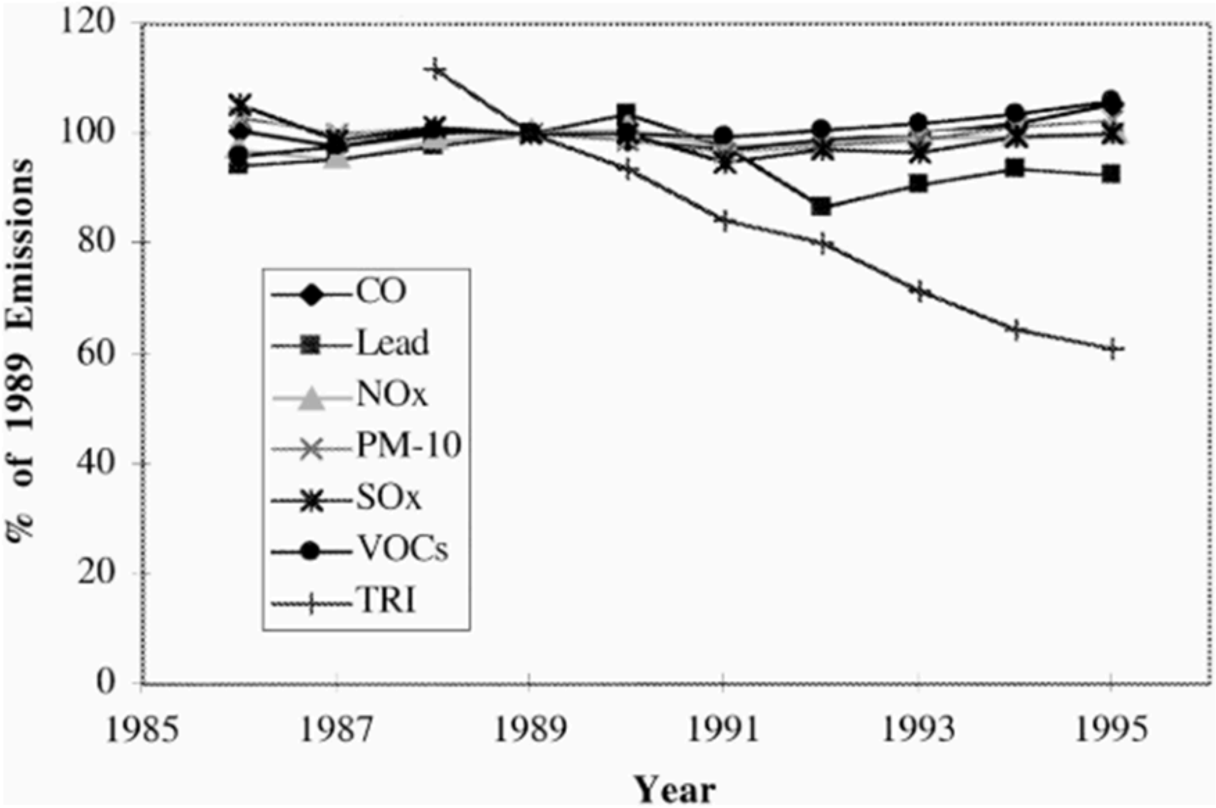
These results are suggestive of the motivations driving *Homo sapiens* to stop and assist strangers in distress. But this leaves the question unanswered as to what might motivate *Homo economicus* to provide assistance. Presumably, *Homo economicus* would be capable of reading the victim's circumstances well enough to make a calculated, rational decision to stop or not, and if having chosen to stop, then how much assistance to provide. Sounds a bit scary if you ask me.

TOXIC RELEASE INVENTORY

In Chapter 8 we considered a game where providing one of two players with additional information about a certain facet of the game ultimately led to a perverse outcome—a reduction in the player's payoff. This result was considered counterintuitive, particularly from the perspective of *Homo economicus* who, according to the rational-choice model, is never supposed to be made worse off when more information is made available. Presuming that *Homo sapiens* could in fact be made better off through the provision of additional information, the US government fostered a natural experiment to test this hypothesis with respect to improving the country's natural environment via the public provision of its data on the levels of toxins emitted by every permitted company in the country. To this day, emissions are self-reported by the polluters, compiled by the Environmental Protection Agency (EPA), and then made publicly available. The EPA inspects approximately 3% of firms per year; one-third of regulated facilities fail to comply with reporting requirements each year.

According to Fung and O'Rourke (2000), between the Toxic Release Inventory's (TRI's) inception in 1988 and 1995, releases of chemicals listed on the TRI declined by 45%. Results are depicted in the figure below where—measured as a percentage of emissions in 1989—the emissions of chemicals included in the TRI diminished steadily over the next five years relative to the emissions of pollutants not included in the TRI.

Those subjects identifying their religiosity as a quest for meaning and who chose to stop also provided a lower level of assistance.



(Fung and O'rourke 2000)

The authors argue that the TRI achieved this regulatory success through the mechanism of "populist regulation," by establishing an information-rich context for private citizens, interest groups, and firms to solve environmental problems. Armed with the TRI, community, environmental, and labor groups can take direct action against the worst polluters, spurring them to adopt more effective environmental practices. Further, the TRI catalyzes popular media campaigns encouraging state-level environmental agencies to enforce regulations against egregious polluters. Additional research also suggests that publicity tied to the government's sharing of TRI data has a negative impact on stock prices of publicly traded firms listed in the TRI. Apparently, *Homo sapiens* from various walks of life are making use of this information to help clean up their local environments; they have been nudged simply via the provision of information in a conveniently accessible database provided by the EPA.

REDUCING DRUNK DRIVING

In 2002, Montana ranked first in the nation in alcohol-related fatalities per vehicle miles traveled. Twenty-one to thirty-year-olds (young adults) represented nearly half of all alcohol-related crashes. Perkins et al. (2010) evaluated the efficacy of a high-intensity, descriptive social-norms marketing media campaign aimed at correcting normative misperceptions about drunk driving, and thereby reducing drinking-and-driving behavior among young adults. Over a 1½ year period, participating counties in the state experienced a "high-dosage" media campaign while non-participating counties experienced a "low-dosage" version.

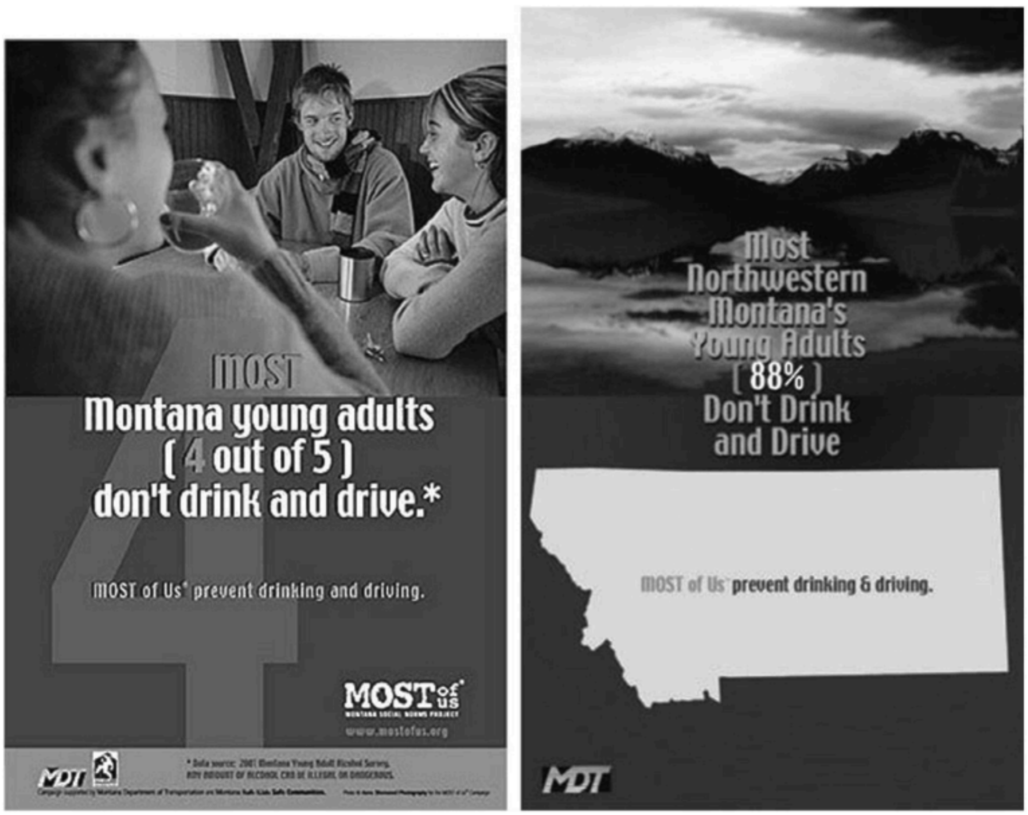
The social norms media campaign consisted of television, radio, print, and theater ads in addition to posters and promotional gifts. Ads appealed to traditional social norms. For example, one television commercial depicted a typical Montana ranch family in a barn preparing to ride horses. As Perkins et al. report, the script read:

> *In Montana, our best defense against drinking and driving is each other. Most of us prevent drinking and driving. We take care of our friends, our families, and ourselves. Four out of five Montana young adults don't drink and drive. Thanks for doing your part.*

Another TV ad depicted a ski lodge window with snow falling. A male voice read the script:

> *In Montana, there are two things you need to know about snow: how to drive on it and how to ski on it. After a day on the slopes and some time in the lodge, my friends and I all take turns being designated drivers." The view widens to reveal the message written on the window, "Most of us (4 out of 5) don't drink and drive." The commercial closes with the voice asking, "How are you getting home?*

One of the posters used in the marketing campaign is depicted below:



(*Perkins et al. 2010*)

The high dosage media campaign ran for 15 months from January 2002 to March 2003. Because many of the intervention counties are sparsely populated (e.g., six are home to fewer than 600 persons in the 21–34-year-old range), Perkins et al. placed a heavy focus on television airtime since not all newspaper and radio advertisements could effectively reach the entire target audience. A total of 18 media advertisements (i.e., 9 television and 9 radio) were used. Social norms advertisements consistently emphasized positive behavior and avoided negative and/or fear-based messages. The television ads were aired during two media flights. The first lasted five and a half months while the second lasted six months. The two radio flights lasted six and a half and six months, respectively.

Television and radio ads were supplemented by local and college newspaper advertisements, theater slides, billboards, various print and promotional items (i.e., t-shirts, key chains, pens, and windshield scrapers), and indoor advertisements. These additional advertisements and theater slides ran from January 2002 through December 2003. Over 250 print ads were taken out in local and college newspapers, 70 theater slides appeared on over twenty movie screens, and a billboard design appeared in seven locations for a two-month period. Over 45,000 promotional items were distributed in the intervention counties. Lastly, 41 indoor ads were placed in Bozeman and Missoula restaurants, which were the two cities with the largest number of individuals from the target population.

The authors measured exposure to the media campaign using both prompted and unprompted recall. Survey participants were asked, "During the last twelve months, do you remember seeing or hearing any alcohol prevention campaign advertisements, posters, radio or TV commercials, or brochures?" If they responded yes, then they were asked what the main message was that they remembered. Participants' perceptions of others' behavior were assessed with two questions: (1) "During the past month, do you think the average Montanan your age has driven within one hour after consuming two or more alcoholic beverages within one hour?", and (2) "In your opinion, among Montanans your age who drink, what percentage almost always make sure they have a designated non-drinking driver with them before they consume any alcohol and will be riding in a car later?"

Lastly, to measure their personal behavior before and after the campaign, survey participants were asked, "During the past month, have you driven within one hour after you have consumed two or more alcoholic beverages within an hour?", and "When you consume alcohol and know that later you will be riding in a car, what percent of the time do you make sure you have a designated non-drinking driver with you before you start drinking?" In addition, participants were asked, "The current law in Montana states that a blood alcohol concentration (BAC) of above 0.10% constitutes legal impairment. Would you support or oppose changing the law in Montana to make a BAC above 0.08% constitute legal impairment? This change would permit less alcohol consumption before driving."

Perkins et al.'s results are presented in the following table:

Differences between intervention and control counties for perceived and reported behavior in November 2001 and June 2003.

	Western intervention counties			Eastern control counties			Intervention–control counties difference of change
	Nov-01	June-03	Change	Nov-01	June-03	Change	Difference
Percent recalling social norms media as main message (unprompted recall)	53.8	70.5	16.7	50.7	42.6	−8.1	24.8***
Percent thinking average Montanan drove within one hour of consuming two drinks in month	91.8	86.7	−5.1	91.9	94.3	2.4	−7.5*
Percent perceiving the majority of peers almost always have a designated driver when drinking and later use car	29.9	39.2	9.3	24.9	23.2	−1.7	11.0†
Percent driving after having two or more drinks within the hour in past month	22.9	20.9	−2.0	16.9	28.6	11.7	−13.7*
Percent reporting they always make sure they have a designated driver when drinking and later use car	41.7	46.4	4.7	42.3	32.0	−10.3	15.0*
Percent supporting changing BAC legal limit for driving to .08	63.5	70.7	7.2	71.1	61.8	−9.3	16.5**

Note. †$p<.06$. *$p<.05$. **$p<.01$. ***$p<.001$.

(Perkins et al. 2010)

The first row of the table presents results for social-norms message recall by participating (or intervention) and non-participating (i.e., control or non-intervention) counties prior to and following the media campaign. As shown in the last column for this row, the campaign was successful at differentially exposing Montanans between the ages of 21 and 34 to social norms messages (the statistically significant difference in message recall across intervention and control counties is 16.7% – (-8.1%) = 24.8%). In the table's second row, we see that the social-norms campaign reduced misperceptions of those in the intervention counties relative to those in the control counties, such that those in the intervention counties believed the average Montanan their same age had driven less often within one hour of consuming two or more drinks in the past month compared to those in the control counties. Similar results were found in the table's third row regarding the perception of peer use of designated drivers. Participants in the intervention counties believed that the majority of Montanans their age almost always have a designated driver with them when they consume alcohol and would be riding in a car later, significantly more so than those in the control counties. As the authors point out, these combined findings suggest that the campaign was successful at reducing normative misperceptions regarding peer drinking and driving behavior.

Relative to participants in the control counties, Perkins et al. find that the percentage of young adults in the intervention counties who reported driving within an hour of consuming two or more drinks in the previous month decreased following the social norms campaign. In contrast, the percentage of young adults in the control counties who reported driving within an hour of consuming two or more drinks in the previous month actually increased during this time. With reported driving after drinking decreasing in the intervention counties by 2% and increasing in the control counties by 12%, there was an overall statistically significant decrease in the intervention counties compared to the control counties of almost 14%.

Similarly, the percentage of individuals in the intervention counties who reported that they always use a designated driver if they plan to drink increased following the social norms campaign, whereas there was a drop in the use of designated drivers in the control counties, resulting in an overall increase in the use of designated drivers in intervention counties relative to the control counties of 15%. Lastly, results indicate that participants in the intervention counties increased their support for changing the BAC legal limit for driving to 0.08 following the social norms campaign, which is a significant difference compared to the decrease in support seen among participants in the control counties. The authors conclude that the social-norms media campaign was effective at reducing high-risk drinking-and-driving behavior and increasing use of protective behaviors (i.e., designated drivers) among those in the intervention counties compared to those in control counties.

If these findings were not enough, Perkins et al. also obtained archival motor vehicle crash records from the intervention and control counties. The authors point out that the data do not provide a perfect test of the intervention's impact because crashes in Montana were coded as alcohol-related when anyone involved in the crash was under the influence of alcohol, regardless of who was driving or at fault. Moreover, the available data only recorded if an alcohol-related crash occurred in the county and not if the driver was from that county. Hence, there may have been some blurring across county lines which, nevertheless, would serve to reduce an observed impact of the intervention (i.e., it would bias the impact of intervention downward). Nevertheless, in spite of these qualifications, the data revealed a pattern in the expected direction. In 2001 there were 9.6% and 10.1% alcohol-related crashes in the intervention and control counties, respectively. However, the difference was not statistically significant. In 2003, after the social-norms media campaign, alcohol-related crashes had declined to 9.1% in the intervention counties and had risen to 10.3% in the control counties, resulting in a statistically significant difference between the two types of counties.

In conclusion, the authors argue that the results of their study provide strong evidence that a comprehensive social-norms media campaign can affect normative perceptions and drinking behavior among young-adult *Homo sapiens*, at least in terms of nudging them away from drinking and driving.

INCREASING VOTER TURNOUT

In a field experiment conducted a month before the 1984 US presidential election, a treatment group of Ohio State University students contacted by telephone was asked to predict whether they would register to vote and whether they would actually vote in the coming days (Greenwald et al., 1987). All students in the treatment group predicted that they would vote, and larger numbers of these students actually registered to vote and voted in comparison with uncontacted students in a control group.

In what Greenwald et al. labeled Experiment 1, a larger percentage of students in the treatment group (who received the simple nudge of a question posed over the phone) registered to vote than did students in the control group (who received no phone call) (20.8% to 9.1%). However, this difference was not statistically significant. To the contrary, the difference in actual voting between the two groups of students (86.7% versus 61.5%) was statistically significant.

In the US, groups like Rock The Vote and Nonprofit VOTE promote a host of different ways to nudge prospective voters to the polls. As Greenwald et al.'s field experiment has shown, it does not really take a big nudge to move the needle on voting. Of course, this is not to say that public calls by some groups to make presidential election day a national holiday, move voting day from the first Tuesday to the first Saturday in November, or move to mail-in voting, are not without merit.

CULTURAL CONFLICT AND MERGER FAILURE

Recall Camerer and Knez's (1994) findings from their laboratory experiments involving mergers in the context of the Weakest Link game. Merged groups obtained inefficient equilibria more frequently than did the separate smaller groups of which the respective merged groups were comprised, suggesting that mergers can exacerbate an extant inefficiency problem in games where inefficient equilibria are focal points, if not the consequence of dominant strategies.

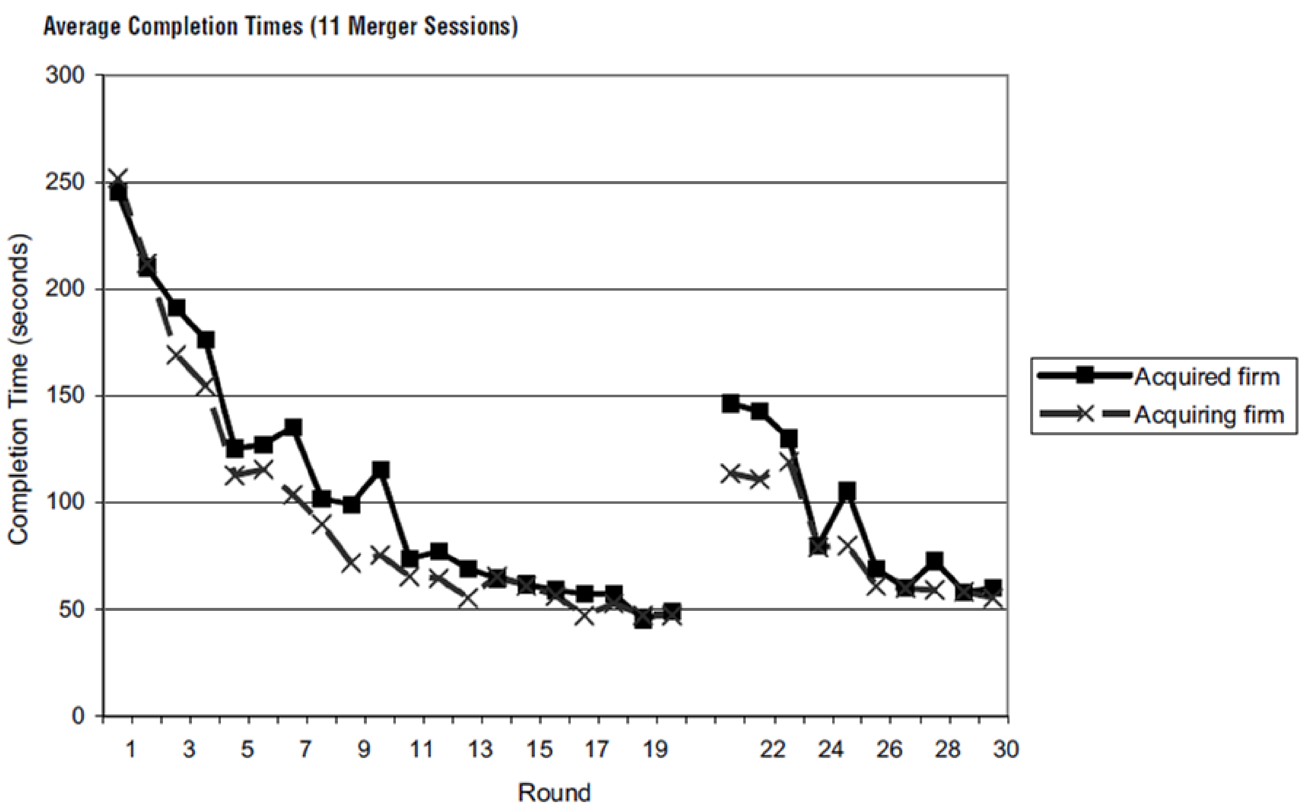
In a novel field experiment, Weber and Camerer (2003) tested for the efficiency effects of mergers by engaging participants in a guessing game similar to Charades. Every subject was shown the same set of 16 photographs, each depicting a different office environment. While most of the photographs shared some common elements (e.g., people, furniture, room characteristics, and so forth), each photograph was unique with respect to the number of people and their characteristics (e.g., gender, clothing, ethnicity), physical aspects of the room (e.g., high ceilings, objects on walls, furniture), and the people's actions (e.g., conversing with others in the picture, talking on the telephone, working at a computer).

Subjects were paired. The experimenter then presented 8 of the 16 photographs in a specific order to only one of the two subjects who had been randomly assigned to play the role of the manager. The manager then described the eight photographs any way she liked to the other subject who, by default, was playing the role of employee. The employee's goal was to select the correct 8 photographs from his collection of the 16 in the same order as presented by the manager as quickly as possible. Each pair of subjects repeated the task for 20 rounds, with the subjects alternating roles of manager and employee each round and the experimenter randomly selecting another set of eight photographs with which to play.

Then two pairs were randomly merged together. One of the pairs was designated "acquiring firm," and one member of this firm was chosen as manager of the now "merged firm" for the remainder of the experiment. The other member of the acquiring firm was designated as an employee, and one member of the "acquired firm" was also selected randomly as an employee. The other member of the acquired firm was now finished participating in the experiment. In the end, therefore, each merged firm consisted of one manager and two employees. Henceforth, the manager played the same game simultaneously with both employees for 10 rounds. However, now each employee completed his or her identification task with the manager independently of the other employee. Each employee tried to guess the eight photographs as quickly as possible. The manager's goal was to achieve the lowest possible average guessing time across the two employees.

As Weber and Camerer point out, the identification task created simple cultures by requiring subjects to develop conversational norms enabling quick reference to the photographs. For instance, one pair of subjects began by referring to a particular picture as "The one with three people: two men and one woman. The woman is sitting on the left. They're all looking at two computers that look like they have some PowerPoint graphs or charts. The two men are wearing ties and the woman has short, blond hair. One guy is pointing at one of the charts" (p. 408). After several rounds, this group's description of the picture had become condensed to simply "PowerPoint."

The study's specific results are presented in the following figure:



(*Weber and Camerer 2003*)

First, note that the average completion time is initially high for the single pairs (what were later to become the acquiring and acquired firms), with both pairs requiring roughly four minutes (250÷60) to complete their tasks in the first round. Average times fell steadily over time as the pairs developed a common language, eventually reaching less than a minute by the 20th round of the experiment. Immediately after the merger (in what is effectively round 21 in the figure), the merged firm performed better than the original pair groups did initially.[43] But relative to where the original pairs

43. In the figure, the merged firm is effectively dissected between (1) the merged-firm's manager paired with the (now former)

ended in round 20 and where the merged firm began, there is a noticeable decrease in the merged firms' performances—from two (acquiring firm) to two and a half (acquired firm) minutes on average for the merged firm compared to less than a minute for the original pair groups. And in the end, the average merged firm never attains the same level of performance as the average original pair.

Surely these results are a rough representation of how mergers actually transpire between acquiring and acquired firms. After all, *Homo sapiens* are a gregarious species. To the extent that management can harness and channel this gregariousness, it could very well be that the merged firm's efficiency is enhanced rather than impeded after the merger (as Weber and Camerer's results suggest). Recall that in the experiment, the employees of the merged firm made their guesses independently of each other. Perhaps if they could have worked together, their overall times would have been reduced sooner and, ultimately (at the end of the tenth round), improved *vis-à-vis* the original pair times. Either way, it is unclear whether *Homo economicus* subjects could have performed any better in this experiment.

MESSY WORK SPACE

In his entertaining book, *Messy: The Power of Disorder to Transform Our Lives*, author Tim Harford explains the multitude of settings in which we *Homo sapiens* are better served by resisting the temptation to tidily organize our lives, particularly in our professional settings (Harford, 2016). Rather, a certain degree of messiness nourishes our creative instincts and often leads to greater productivity. Harford provides several examples, from musicians Brian Eno's, David Bowie's, Miles Davis, and Keith Jarrett's "messy" abilities to improvise and embrace randomness; to Martin Luther King Jr.'s willingness to go off-script in his most famous "I Have a Dream" speech; to the spontaneous, self-deprecating humor of Zappos customer service reps; to the daring battlefield maneuvers of General Rommel; to Amazon founder Jeff Bezos' counterintuitive business strategy; to world chess champion Magnus Carlsen's puzzling chess moves; to the nomadic, freewheeling methods of scientists Erez Aiden and Paul Erdős, and the eclectic and messy collaborations among the scientists inhabiting Building 20 on the campus of the Massachusetts Institute of Technology (MIT) during the last half of the 20th century. In each case, it was the messiness of the individual's or group's method that helped empower their sense of self and, ultimately, motivate their success.

In one particular study cited by Harford, authors Knight and Haslem (2010) designed a field experiment to measure the extent to which office workers' freedom to customize, or mess, their workspaces—and thus, empower themselves—impacts their wellbeing (i.e., feelings of psychological comfort, organizational identification, physical comfort, and job satisfaction, as well as their overall productivity).[44] As the authors point out, companies have traditionally believed that lean, open, uncluttered office spaces are efficient. These types of spaces can accommodate more people and thus, exploit economies of scale. Desks can also easily be reconfigured for use by other workers. As a result, space occupancy can be centrally managed with minimal disruptive interference from workers. Surprisingly, there is a lack of empirical evidence to support these claims.

To the contrary, some organizations have sought to enrich their workspaces by investing in "environmental comfort" (e.g., aesthetically pleasing artwork and living plants) to enhance the physical

employee of the (former) acquiring firm—denoted by the Xs - and (2) the merged-firm manager paired with the (now former) employee of the (former) acquired firm—denoted by the black squares.

44. In a similarly interesting study of individual work habits, Whittaker et al. (2011) find that messy email sorters (i.e., office workers who do not organize their saved email messages into different folders) are generally more efficient time-wise when it comes to retrieving the saved messages for current use.

and mental health of their employees. And some of these organizations go even further by encouraging employees to decorate their personal workspaces with meaningful artifacts to project their identity onto their environment and to give some sense of permanency, control, and privacy. In this respect, these organizations encourage their employees to messy their workspaces.

Given this dichotomy between "lean" and "enriched" workspaces, Knight and Haslem designed a field experiment to directly test the hypotheses that empowering workers to manage, and have input into, the design of their workspace enhances their sense of organizational identification, emotional well-being, and productivity. Four separate office spaces were designed for the experiment: (1) a lean, minimalist office space intended to focus the employee's attention solely on the work at hand (in particular, through the imposition of a clean desk policy), (2) an enriched office space incorporating art and plants, but where the employee has no input into their arrangement, (3) an empowered office space that allows an employee to design her office environment using a selection of the same art and plants as in the enriched office space, but allows her to realize something of her own identity within the working space, and (4) a disempowered office space where the employee's workspace design in the empowered office is overridden by the experimenter so that an initial sense of autonomy within the workspace is effectively revoked.

In the experiment, 112 men and women ranging in age from 18 to 78 years were randomly assigned to one of the four office space types.[45] The laboratory office was a small interior office measuring 3.5m x 2m. The office had no windows or natural light. At the outset of the experiment, each participant was left alone in the office space for five minutes to take in the ambient environment. The office contained a rectangular desk and a comfortable office chair. The room was lit by diffused, overhead fluorescent tubes, the floor was carpeted, and an air conditioning system kept the room at a constant temperature of 21 °C.

In the lean office, no further additions to the room were made. In the enriched office, six potted plants had already been placed toward the edge of the desk surface, so as not to impinge on the participants' working area. Six pictures were hung around the walls. The pictures were all photographs of plants enlarged onto canvas. In the empowered office, the pictures and plants had been placed randomly around the room. Participants were told that they could decorate the space to their taste using as many, or as few, of the plants and pictures provided. They could, therefore, work in a lean or very enriched space or at a point anywhere along that continuum. The disempowered office was the same as the empowered office; however, when the experimenter re-entered the office, she looked at the chosen decorations, briefly thanked the participant, and then completely rearranged the pictures and plants, thereby overriding the participant's choices. If challenged, participants were told that their designs were not in line with those required by the experiment.

After getting situated in their offices, participants were instructed to perform a card-sorting task. Three packs of playing cards had been shuffled together, and the participant was required to sort them back into the three constituent packs and then to sort each pack into its four suits (hearts, clubs, diamonds, and spades). These suits then had to be ordered from ace to king and placed in discrete piles, leaving 12 piles in total. The key performance measures were the time taken to complete this task and the number of errors made.

The participants were then asked to perform a second "vigilance task," whereby they were given a photocopy of a magazine article and asked to cross out and count all the lowercase letters "b"

45. The authors actually conducted two separate experiments with different samples of participants and a slightly adjusted experimental design. The results of these experiments were consistent with those of the experiment reported here.

that were on the page. The time taken to complete the task was measured as well as the number of errors (missed b's). After completing this task, the participants completed a 74-item questionnaire, answering questions that would enable the researchers to test the previously mentioned hypotheses.

Knight and Haslem found that the average participant performed the card sorting task best in the empowered office space. When measured in minutes to complete the task, the results (i.e., the measured productivity differences between the lean, enriched, empowered, and disempowered offices) are statistically different for two of the study's three main hypotheses – the first two hypotheses (H1 and H2, respectively) that the empowered office inspires the highest productivity level (particularly when compared with the enriched office) are confirmed. Regarding the vigilance task (measured in minutes to complete the task), the average participant again performs best in the empowered office; interestingly, the disempowered office space is associated with a particularly deleterious effect on productivity. In terms of total productivity measured in minutes to completion across the two tasks, the empowered office environment inspires the most productivity, while the disempowered office space inspires less productivity vis-à-vis the enriched office space.

With respect to the more intangible effects on well-being and organizational identification, the typical participant's sense of involvement, autonomy, and psychological comfort ranked highest in the empowered office space, particularly in the enriched office space. Regarding one's sense of identification with the organization (in this case, the field experiment's various tasks), the only statistically significant effect occurred in the disempowered office. Disempowerment resulted in a decrease in organizational identification.

The message from Knight and Haslem's field experiment seems clear. Empowering office workers, which, from the perspective of managers, risks introducing a degree of messiness into workspaces, can result in greater productivity and a sense of well-being among employees. Empowerment does come with a risk though. Once the empowerment genie is out of the proverbial bottle, woe to the manager who tries to stuff it back in.

MESSY TRAFFIC CROSSING

Given *Homo sapiens*' predisposition for all things tidy, one can be forgiven for concluding that Dutch traffic engineer Hans Monderman had lost his mind when he argued that traditional traffic-safety infrastructure—warning signs, traffic lights, metal railings, curbs, painted lines, speed bumps, etc.—is often unnecessary and, worse, can endanger those it is meant to protect. Yet according to Vanderbuilt (2008), this was indeed Monderman's sentiment, based as it was upon the Dutch traffic guru's simple axiom, 'when you treat people like idiots, they'll behave like idiots.' To wit, Monderman devoted the better part of his career designing roads to feel more dangerous (yes, "more") so that pedestrians and drivers would navigate them with greater care (Vanderbuilt, 2008).[46]

As recounted by Vanderbuilt, Monderman's most memorable design was built in the provincial Dutch city of Drachten in 2001. At the town center, in a crowded four-way intersection called the *Laweiplein*, Monderman removed not only the traffic lights but virtually every other traffic control. Instead of a space cluttered with poles, lights, "traffic islands," and restrictive arrows, Monderman installed a radical kind of roundabout (which he called a squareabout because it resembled more a

46. A similar type of effect is exhibited by children at play in what we would commonly agree are less-safe, more-risky playground environments. As Harford (2016) points out, children naturally adjust for risk—if the ground is harder, the play equipment sharp-edged, the spaces and structures uneven, they choose to be more careful. Learning to be alert to risk better prepares the children for self-preservation in other settings.

town square than a traditional roundabout) marked only by a raised circle of grass in the middle, several fountains, and some very discreet indicators of the direction of traffic, which were required by law. Rather than creating clarity and segregation, Monderman had created confusion and ambiguity in the minds of drivers and pedestrians. Unsure of what space belonged to them, drivers became more accommodating and communicative. Rather than give drivers a simple behavioral mandate—say, a speed limit sign or a speed bump—Monderman's radical design subtly suggested the proper courses of action.

A year after its redesign, the results of this extreme makeover were striking. According to Euser (2006), not only had congestion decreased in the intersection—buses spent less time waiting to get through the intersection, for example—but there were half as many accidents even though total car traffic had increased by a third. Further, both drivers and, unusually, cyclists were signaling more often. Despite the measurable increase in safety, local residents perceived the squareabout to be more dangerous.

Roughly five years after the redesign, Euser found that on the busiest street entering into the squareabout, the average waiting times for automobiles had dropped from 50 to about 30 seconds. Waiting times for public buses dropped from over 50 seconds to 26 seconds heading in one direction and to 38 seconds heading in the other. The number of cyclists entering the intersection between 3:30 pm and 5:30 pm on a typical day and signaling with their left hands increased from roughly 50% to 80%, while the percentage of right-hand signaling increased from 9% to 47%. Underscoring these changes in waiting times and cyclist behaviors, the number of cyclists accessing the intersection in the same two-hour timeframe on a typical day had increased by roughly 5% since 2000. The number of "person car units" (PCUs) entering the intersection during the typical evening rush hours had increased by roughly 30%. Three years after the squareabout's construction, the total number of traffic accidents had been roughly cut in half.

In surveys conducted with Drachten residents both before and after construction of the squareabout, traffic was generally considered less safe, particularly among the elderly. Motorists and cyclists reported feeling less safe, while there was no discernable change in the perceived safety of pedestrians. In terms of the *Laweiplein*'s spatial quality, survey respondents generally reported experiencing an improvement. In general, bus drivers reported feeling positive about the new design, in particular, that their wait times had improved.

Though Monderman's squareabout may be counterintuitive and revolutionary, clearly the change in *Homo sapiens*' traffic behaviors it has impelled is yet another indication of *Homo sapiens*' susceptibility to nudges which, in this case, involves the redesign of existing infrastructure.[47] As for *Homo economicus*, whose rational mind is enthralled by tidiness and order, navigating the squareabout is something she would never allow herself to get used to.

DISORGANIZED PEDESTRIANS

I don't know about you, but in my daily life I suffer from what might be called myriad "micro-inefficiencies." I waste minutes each day searching for things—my cellphone charging cable, my glasses, keys. And then, I fumble with these things once I have found them—extricating the charging

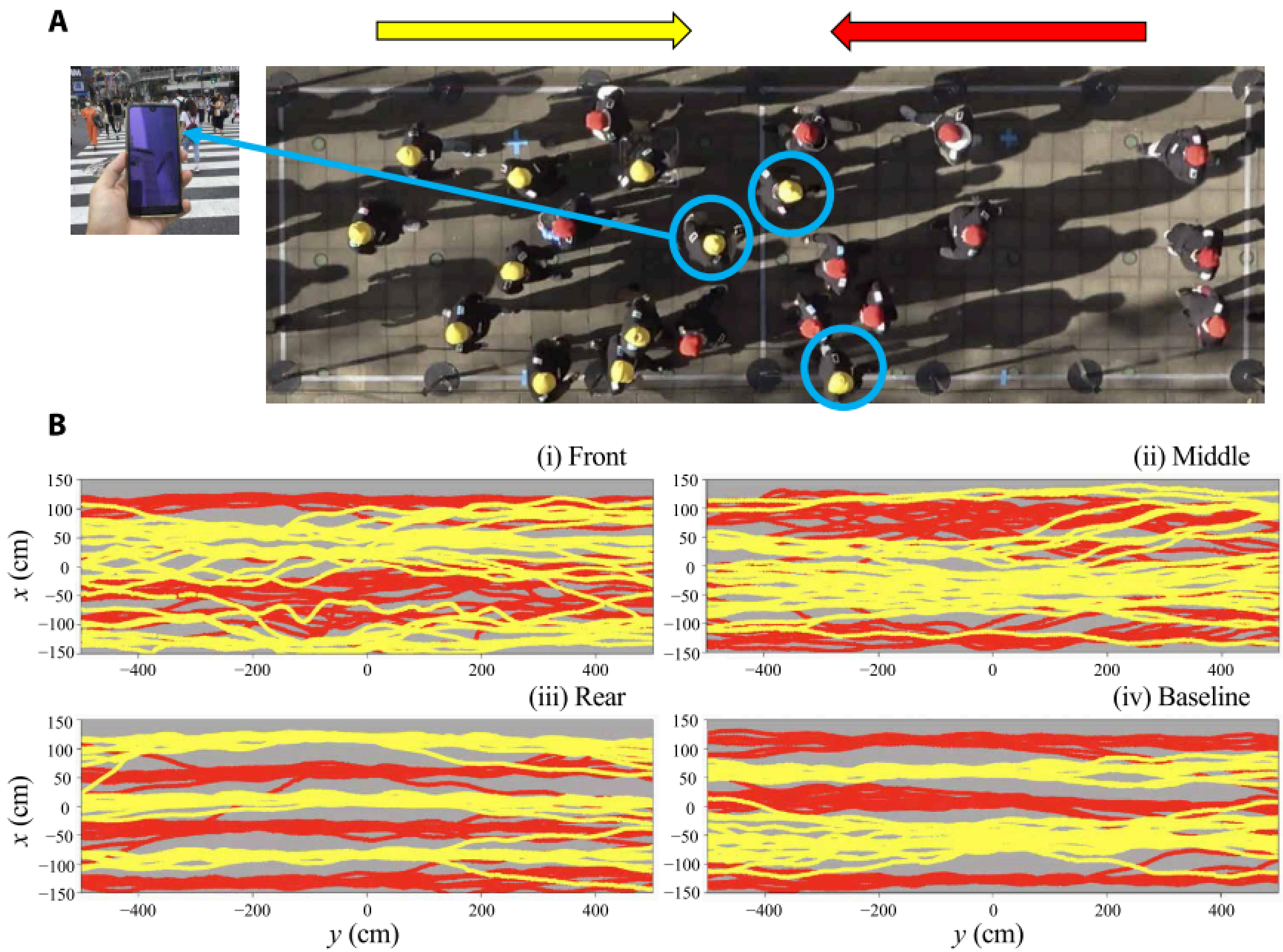
47. Interestingly enough, my traffic experiences in the Southeast Asian nation of Myanmar, where the seed for writing this book was planted, suggest that in the absence of infrastructure, motorists (both in cars and on scooters) can reach an equilibrium surprisingly free of major accidents with the repetitive and sometimes symphonic use of their horns, well-honed intuition, and what appears to be highly developed peripheral vision.

cable from its little stuff sack, fishing my keys out of my pocket, taking my glasses off to see something up close and then forgetting where I left them. These micro-inefficiencies add up over a lifetime, perhaps claiming a year or more off my lifespan.

Living in a small town, I am also fortunate to benefit from certain "micro-efficiencies." No matter where I travel in town or the surrounding valley, I rarely encounter congestion, whether on the road walking, riding my bicycle, driving my car, walking or jogging on a sidewalk, or waiting in line at a restaurant or the grocery store checkout. So, in many respects, I should not complain about my micro-inefficiencies. This is especially the case when I travel to busier cities—the traffic congestion in Salt Lake City and Seattle never ceases to amaze and humble (okay, and often frustrate) me. I also experience similar feelings of amazement, humility, and frustration when walking the bustling New York City sidewalks in midtown Manhattan in and around Grand Central Station.

As a visitor to the city, I am inclined to gawk at the buildings, the street life, and the captivating mix of other pedestrians. At the same time, I am tasked with having to navigate the sidewalks without bumping into or impeding the flow of other pedestrians. The flow of pedestrians in this section of New York City might best be described as organized disorganization. It is as if we *Homo sapiens* tacitly and instinctively coordinate to reduce what would otherwise result in more substantial micro-inefficiencies in our lives. Like fish schooling in the oceans and birds flocking in the sky, New York City pedestrians self-organize to reduce the incidence of collision. As Murakami et al. (2021) point out, pedestrians' instantaneous decisions are influenced more by anticipated future positions rather than the current positions of their nearest neighbors, which suggests that crowded pedestrians are not just passively repelled by other pedestrians, but actively discern passages through a crowd by anticipating and tacitly negotiating with neighbors to avoid collisions in advance.

To test the sensitivity of these tacit negotiations, Murakami et al. conducted a simple field experiment of lane formation, where some participants walked while using their cellphones, thus potentially interfering with their ability to anticipate neighbors' motions. Two groups of 27 pedestrians each voluntarily agreed to walk in bidirectional flows in a straight mock corridor. Three participants in one of the two groups were visually distracted by using their cellphones to potentially disrupt their anticipatory interactions with the other pedestrians in both their group and the other group. This situation is depicted in panel A of the figure below. The three circled individuals with yellow hats are looking at their cellphones as they move with the other yellow-hatted pedestrians from left-to-right in the corridor. The red-hatted pedestrians are all moving right-to-left.



(Murakami et al. 2021)

The authors hypothesized that distracted pedestrians located in the front of their group, directly facing the oncoming crowd, would have the most influence on overall crowd dynamics. To test this, Murakami et al. designed three treatment scenarios. In one treatment, the three randomly selected cellphone users were positioned at the front of their group (front treatment), while in the other two treatments three randomly selected cellphone users were placed in the middle (middle treatment) and rear (rear treatment), respectively, of their group. In a control scenario, no one was selected to use their cellphone. The experiments were replicated 12 times for each of the treatment and control scenarios.

The authors found that pedestrians participating in the front treatment were significantly slower than those participating in the control scenario, suggesting that distracted participants in the front condition influenced overall pedestrian flow as expected (i.e., the three distracted pedestrians created a micro-inefficiency for the other pedestrians). This inefficiency is depicted by the preponderance of red and yellow overlapping squiggly lines in the above figure's graph (i) of panel B relative to the absence of overlapping squiggly lines in graph (iv). However, pedestrians participating in the middle and rear treatments were not found to be significantly slower than participants in the control scenario (depicted by comparing graphs (ii) and (iii) with graph (iv) in the figure).

Just for fun, check out these videos of the experiment. The first video is of the control scenario, where no pedestrians are looking at their cell phones:

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://uen.pressbooks.pub/behavioraleconomics/?p=1182#video-1182-1

The second video shows the front treatment, where the three cell phone users are located at the front of the red-hatted group of pedestrians walking left-to-right:

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://uen.pressbooks.pub/behavioraleconomics/?p=1182#video-1182-2

The third and fourth videos are of the middle and rear treatments. In the middle treatment, the three cellphone users are located in the middle of the yellow-hatted group walking right-to-left, and in the rear treatment, the three cellphone users are located at the rear of the red-hatted group walking left-to-right.

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://uen.pressbooks.pub/behavioraleconomics/?p=1182#video-1182-3

One or more interactive elements has been excluded from this version of the text. You can view them online here: https://uen.pressbooks.pub/behavioraleconomics/?p=1182#video-1182-4

BENEFICIAL BIASES IN STRATEGIC DECISION-MAKING

Are *Homo sapiens* who start their own small businesses (i.e., entrepreneurs) fundamentally different from those who choose to work in larger, more-established businesses (i.e., managers)? This question has spurred a long line of research concerning the mindset and behavior of entrepreneurs, with findings suggesting that entrepreneurs are risk-seekers and rugged individualists (e.g., McGrath et al., 1992), social deviates (Shapero, 1975), and a breed apart (Ginsberg and Buchholtz, 1989). In their survey of entrepreneurs and managers, Busenitz and Barney (1997) probe why the decision-making processes of these two types of *Homo sapiens* vary regarding how they manifest well-known biases.

The authors find that entrepreneurs fall prey to biases to a greater extent than managers, in particular biases associated with optimistic overconfidence and representativeness.[48]

As Busenitz and Barney point out, overconfidence tends to manifest itself more in entrepreneurs' decision-making processes, which enables them to proceed with ideas before each step in a venture is fully known. In the face of uncertainty, a higher confidence level can encourage an entrepreneur to take an action before it makes complete sense to do so. Representativeness also manifests itself in entrepreneurs' decision-making processes via a propensity to short-cut by generalizing results from small, nonrandom samples such as their personal experiences with customers.

Busenitz and Barney also point out that respective decision-making contexts further distinguish entrepreneurs from managers. On average, decisions facing entrepreneurs are made in more uncertain and complex environments. Large organizations develop extensive policies and procedures to aid and inform managers; managers usually have access to information on historical trends, past performance, and other market-based information. To the contrary, entrepreneurs rely instead upon simplifying biases and heuristics to exploit brief windows of opportunity.

The authors' sample of 176 entrepreneurs was drawn from plastic manufacturing, electronics, and instruments—more dynamic industries representing a higher percentage of newly emerging firms. To be considered an entrepreneur, a survey respondent had to have been a founder of a firm and had to be currently involved in the firm's start-up process—criteria that reduced the entrepreneur sample down to 124. For the manager sample, managers in large organizations had to have responsibility for at least two functional areas (e.g., marketing and finance, personnel and research and development) and work for a publicly owned firm with more than 10,000 employees.

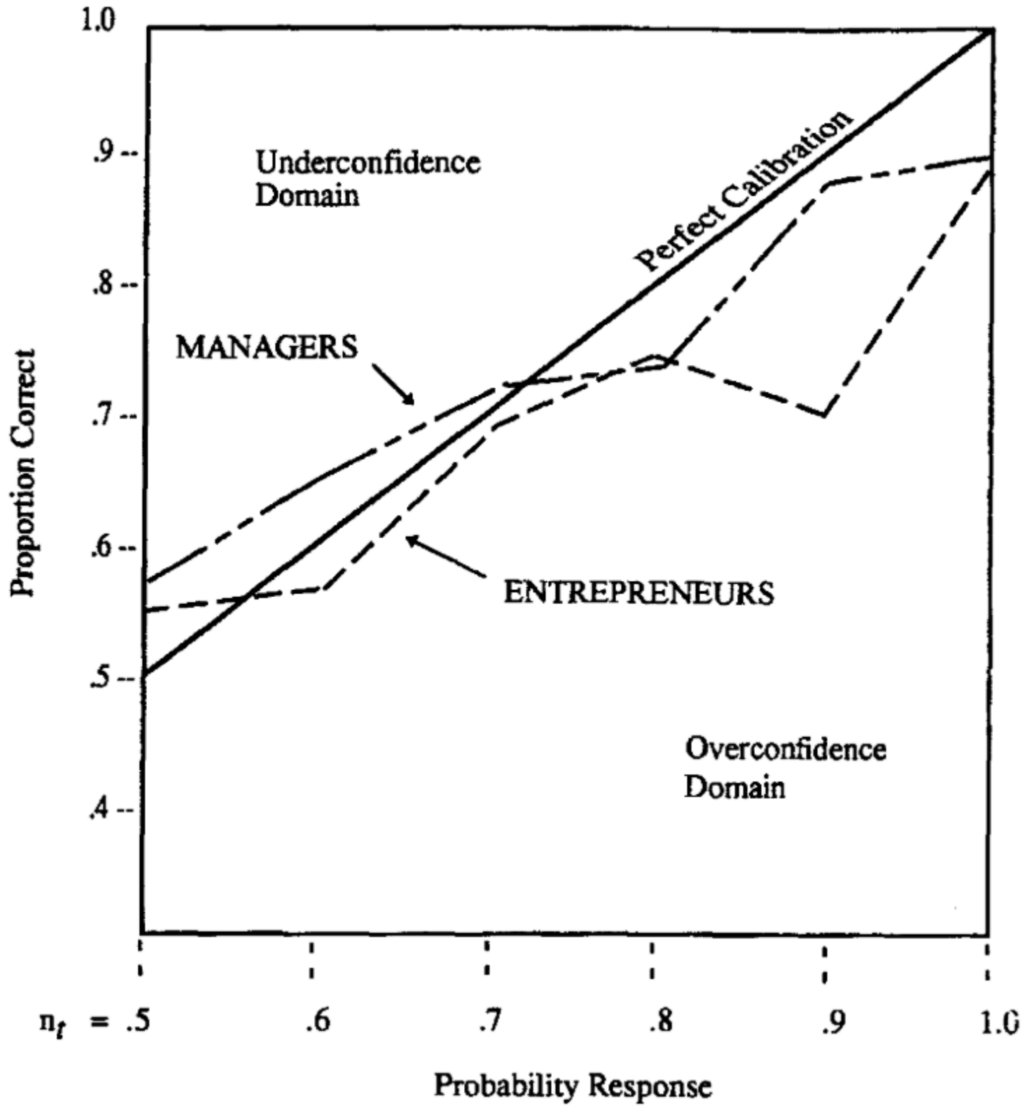
To measure overconfidence, subjects were presented with a series of five questions concerning death rates from various diseases and accidents in the US. For example, one question was, Which cause of death is more frequent in the US, cancer of all types or heart disease?[49] Subjects provided two responses to each question. First, they chose one of the two alternatives as their best guess of the correct answer. Second, they stated their confidence in their choice based upon a scale ranging from 50% to 100% confidence, where 50% indicated their answer was a total guess, and, say, 70% indicated that they had a 70% probability of being correct. A statement of 100% indicated that a subject was certain their answer was correct. A summary measure of overconfidence was then calculated where a positive score represented overconfidence and a negative score under-confidence. As an example of how the score was derived, suppose a subject's confidence statements for the five questions were 50%, 60%, 70%, 70%, and 90%, respectively, resulting in a mean confidence percentage of 68%. Further suppose the subject provided correct answers for three out of the five questions (i.e., for 60% of the questions). This subject's confidence score would then be calculated as 68% – 60% = 8%.

The figure below represents the aggregate results for correct responses by confidence category (50%, 60%, 100%) divided by the total number of responses across subjects in each category. For example, if all the responses in the 70% confidence category for managers were correct 70% of the time, then the grouping at the 70% confidence level is perfectly calibrated. The reference line labeled "perfect calibration" depicts perfect calibration at each respective confidence level. The two lines labeled "managers" and "entrepreneurs" indicate that entrepreneurs are overconfident in their choices

48. Recall that overconfidence results when an individual is overly optimistic in his initial assessment of a situation and then too slow in incorporating additional information in his reassessment. Representativeness results when individuals willfully generalize about phenomena based upon only a few observations (recall the base-rate example from Chapter 2).

49. Correct answers were based on information provided by the National Center for Health Statistics.

at five of the six confidence levels, whereas managers are overconfident at only three out of the six levels. We also see that entrepreneurs are more overconfident than managers at each of the confidence levels except at 80% where the groups are nearly identical.



(Busenitz and Barney 1997)

To measure representativeness, Busenitz and Barney provided subjects with two separate scenarios representing various types of real-life strategic decisions. Each scenario consisted of two alternatives, one of which subjects chose as their preferred alternative. Scenario 1 involved the purchase of a major piece of equipment, whereas scenario 2 depicted an automation update decision. After deciding on their preferred alternative, subjects described their reasoning behind each decision. Coders then analyzed the responses to determine the extent to which heuristic-type reasoning was used by subjects in determining their preferred alternative. A code of "1" was assigned to responses that contained no mention of statistical reasoning, relying instead on subjective opinions or simple rules of thumb. Examples of this form of reasoning include reference to personal experience or simple decision rules like "buy American." A code of "0" was assigned to responses containing some form of statistical reasoning, including references to variability or sample size. Finally, the results for both scenarios were summed to create a single three-category variable (0-2), with a "0" indicating that the subject used statistical reasoning across both problems, and a "2" indicating that only heuristic reasoning was used.

The table below presents results for logistic regression analysis, where, for those of you familiar with this type of analysis, the dependent variable represents entrepreneur versus manager (coded "2" if the former and "1" if the latter). Given this coding of the model's dependent variable as well as the coding for the overconfidence and representative measures, we expect coefficients associated with these measures to be positive and statistically significant, which, as we see, they are.

	Model 1		Model 2	
Independent Variables	Parameter Estimate	Wald χ^2	Parameter Estimate	Wald χ^2
Intercept	2.07^c	30.64	6.31^c	17.6
Risk-Taking			−0.007	0.005
Conformity			-0.39^b	11.64
Education			-1.09^c	31.76
Age			−0.03	1.67
Alertness			0.34^a	6.61
Representativeness	1.6^c	36.49	1.56^c	22.09
Overconfidence	2.68^b	8.36	2.76^a	6.21
Pseudo-R^2	0.21		0.37	
Model χ^2	54.43^c		108.5^c	
df	198		185	
Hit ratio (%)	70		79^c	

$^a p < .05.$
$^b p < .01.$
$^c p < .001.$

(Busenitz and Barney 1997)

Model 1 includes solely the representativeness and overconfidence measures as explanatory variables, whereas Model 2 includes a host of control variables measuring a subject's proclivity for risk-taking and conformity, degree of alertness, as well as level of education and age. In both models, the coefficients for representativeness and overconfidence are indeed positive and statistically significant, and these two measures by themselves correctly distinguish entrepreneurs from managers more than 70% of the time (as indicated by the Hit Ratio statistic). These results suggest that Representativeness and Overconfidence Biases manifest themselves more in the strategic decision-making behavior of entrepreneurs than managers.

It should come as no surprise that differences such as these would not arise with *Homo economicus* entrepreneurs and managers since neither group would be susceptible to Representativeness and Overconfidence Biases to begin with.

BENEFICIAL HEURISTICS TOO

The implication of the two heuristics discussed in Chapter 1—the Affect and Availability Heuristics—is that heuristics generally lead to misjudgments on the part of *Homo sapiens*. As Gigerenzer and Gaissmaier (2011) point out though, this is not always the case. The use of heuristics can sometimes lead to more preferable outcomes than those determined by statistical analyses or more complex strategies.

For example, consider the Hiatus Heuristic, where a customer who has not purchased anything from a business within a certain number of months (the hiatus) is classified as inactive. In their study of an apparel retailer, an airline, and an online CD retailer—whose Hiatus Heuristics were nine, nine, and six months, respectively—researchers Wübben and Wangenheim (2008) compared this heuristic to a statistical analysis of 40 weeks of data from each company.[50] The heuristic resulted in a correct classification of an inactive customer if the customer did not make a purchase during the 40-week period of analysis (which is one month longer than the nine-month hiatus period for the apparel retailer and airline, and four months longer than the CD retailer's six-month hiatus period).[51]

For the apparel retailer, the Hiatus Heuristic correctly classified 83% of customers, whereas the statistical model classified only 75% correctly. For the airline, the Hiatus Heuristic correctly classified 77% of the customers versus 74% for the statistical model, and for the CD retailer, the two approaches each correctly classified 77% of the customers.

Another beneficial heuristic identified by Gigerenzer and Gaissmaier (2011) is known as the "Recognition Heuristic," whereby if an individual first determines a criterion upon which to judge the value of some alternative, say alternative A, and alternative A is later recognized among the other alternatives, then the individual infers that alternative A has the higher value associated with that criterion.[52] For example, in predicting federal and state election outcomes in Germany, forecasts based upon surveys of how well voters recognized the candidates' names performed almost as well as interviews with voters about their actual voting intentions (Gaissmaier and Marewski, 2010). Similarly, in three studies conducted by Ortmann et al. (2008) designed to predict the stock market, recognition-based portfolios (i.e., stock portfolios comprised of the most-recognized companies) on average outperformed managed funds such as the Fidelity Growth Fund, the market (Dow or DAX), chance portfolios, and stock experts.

The One-Clever-Clue Heuristic is used in a myriad of circumstances. For example, Snook et al. (2005) study the use of geographic profiling to predict where a serial criminal is most likely to live given the locations of a series of crimes. Typically, geographical profiling utilizes a sophisticated statistical software program to calculate probability distributions across possible locations. The authors tested a special case of the One-Clever-Clue Heuristic, which they named the Circle Heuristic. The Circle Heuristic predicts the criminal's most likely location simply as the center of a circle drawn through the two most distant crime locations. The heuristic thus relies on one cue only: the largest distance. In a comparison with 10 other profiling distributions, the Circle Heuristic predicted the locations best. Nevertheless, the authors found that complex profiling strategies became more accurate as the number of crime locations was no less than nine.

As Gigerenzer and Gaissmaier (2011) point out, the Take-The-Best Heuristic is similar to the One-Clever-Clue Heuristic, except that clues are retrieved from the individual's memory. Perhaps the most famous example of how the Take-The-Best Heuristic has been used effectively is Green and Mehr's

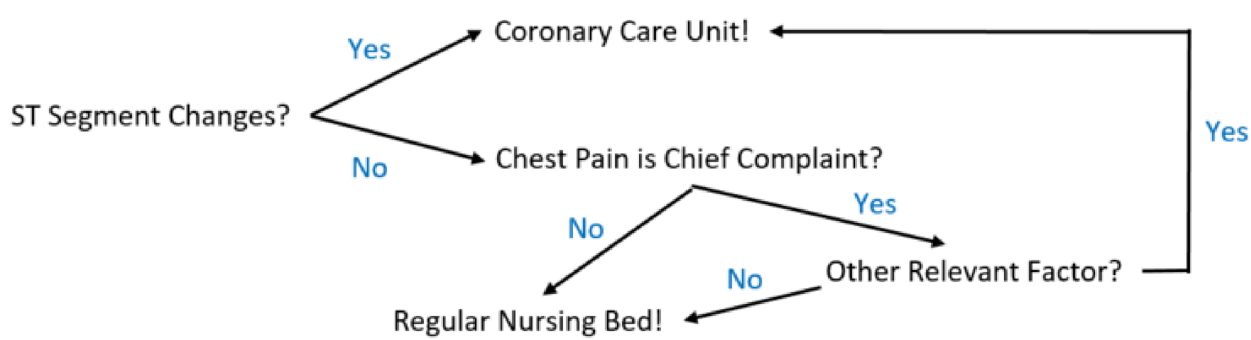
50. For those of you with a background in econometrics, the authors estimated the parameters of a negative binomial model where the count variable is the number of purchases made during the respective hiatus periods.

51. The statistical model classified a customer as being inactive if the model's prediction was that the customer did not make a purchase during the 40-week period.

52. A related heuristic, known as the Fluency Heuristic, is used when alternative A is merely recognized more quickly than the other alternatives.

(1997) study of how patients are assigned to the coronary care unit (CCU) at a Michigan hospital. When a patient arrives at any hospital with severe chest pain, emergency physicians have to decide quickly whether the patient suffers from acute heart disease and should therefore be assigned to the CCU. Historically, at the Michigan hospital in question doctors preferred to err on what they considered to be the safe side by sending 90% of the patients to the CCU, although only about 25% of these patients typically had critical symptoms. This created overcrowding in the CCU and a concomitant decrease in quality of care.

Green and Mehr (1997) tested two approaches. One was the use of a logistic regression equation known as the Heart Disease Predictive Instrument (HDPI). The other was a simple decision tree depicted in the figure below. Accordingly, if the answer to the question of whether an ST segment has changed is "yes," then the patient is immediately sent to the CCU. If "no," and the patient's chief complaint is chest pain, his condition is assessed for one of a few other factors. If a factor is present, the patient is then sent to the CCU. If the patient's ST segment has not changed and he either does not complain of chest pain or another factor is not present, then the patient is provided with a regular nursing bed.



To use the HDPI, doctors review a complex chart consisting of approximately 50 different possible symptoms and enter what they consider to be the probabilities of relevant symptoms into a pocket calculator. Green and Mehr (1997) found that the decision tree was more accurate in predicting actual heart attacks than the HDPI—the use of the decision tree sent fewer patients who suffered from a heart attack to a regular bed and also nearly halved physicians' high false-alarm rates (i.e., the sending of patients to the CCU when they instead should have been sent to a regular bed). Because the decision tree was a transparent and easy-to-memorize heuristic, the hospital's physicians preferred its use.

As this brief discussion indicates, there are a plethora of ways in which *Homo sapiens* actually outsmart *Homo economicus* through the use of heuristics. Indeed, Gigerenzer and Gaissmaier (2011) report on a study showing that a simple physician's bedside exam can outperform the use of a magnetic resonance imaging (MRI) exam in the diagnosis of a brainstem or cerebellar stroke, and another showing how the simple 1/N Rule, where resources (e.g., time or money) are allocated equally over N different alternatives, can outperform more sophisticated optimization models. These are specific examples of Trade-Off Heuristics. The point is that the simpler mind of *Homo sapiens* does not always put her at a disadvantage relative to the more complex mind of her *Homo economicus* muse.

IF YOU GIVE A GROCERY SHOPPER A MUFFIN

Gilbert et al. (2002) ran a field experiment to test for the presence of Projection Bias among roughly 100 grocery store shoppers. The shoppers were stopped in the parking lot on their way into the store and asked to participate in a Taste Test and Survey Study. Participating shoppers began by making a list of the items they planned to purchase that day in the store (shoppers who had already created their own lists were politely informed that they were ineligible to participate). Participants were randomly chosen to either receive their lists back before entering the store (henceforth known as "listful" shoppers), or not receive their lists back ("listless" shoppers). Next, participants were assigned to one of two groups. One group was given muffins to eat before entering the store (henceforth known as "sated" shoppers). The other group was asked to return after having completed their shopping to pick up their muffins ("hungry" shoppers). On average, sated shoppers both listful and listless later self-reported themselves as feeling less hungry after having completed their shopping trip than listless and listful hungry shoppers.

After shopping, each shopper's receipt was collected. The authors found that, on average, listless sated shoppers made significantly fewer unplanned purchases than listless hungry shoppers. In specific, over 50% of the purchases made by the listless hungry shoppers were unplanned, versus only 34% of purchases made by listless sated shoppers. The difference between listful sated and listful hungry shoppers was not statistically significant. The authors concluded that when we are hungry, food is attractive. When unconstrained by a pre-made shopping list, the items we pass in the grocery store are thus more likely to be evaluated based upon how they would satisfy our current, rather than future, hunger.[53] As such, listless sated shoppers effectively de-contaminate their hedonic mental representations of consuming their food purchases in the future by shopping with less hunger in the present (i.e., they are less likely to suffer from Projection Bias of their future food consumption needs). Shopping with the intent of accurately satiating future hunger is, after all, the point of one's weekly grocery trips!

CATALOG SALES AND PROJECTION BIAS

It is no surprise that in the rational-choice model of *Homo economicus*, individuals accurately predict how their tastes for different goods and services change over time. It should likewise come as no surprise that no such accuracy attends the predictions made by *Homo sapiens*. Loewenstein et al. (2003) hypothesized that *Homo sapiens* exhibit a systematic Projection Bias when it comes to accounting for changes in their tastes. While they tend to understand the direction in which their tastes change (e.g., that eating the same foods for dinner over and over diminishes one's appetite for those foods), *Homo sapiens* systematically underestimate the magnitudes of these taste changes.

In a nifty field experiment, Conlin et al. (2007) set out to test this hypothesis by analyzing catalog orders for weather-related clothing items and sports equipment. The authors find evidence of Projection Bias with respect to the weather, in particular, that *Homo sapiens* are overinfluenced by weather conditions at the time they make decisions on what to order. Specifically, if the weather on the day he places an order from the catalog would make the item seem more valuable if used on that day, then he is more prone to order the item. For example, with cold-weather items (i.e., items that

53. Loewenstein (2005) would say that, upon entering the grocery store, hungry shoppers were in a "hot state" of mind while sated shoppers were in a "cold state." The difference between these two states of mind is what Loewenstein (2005) calls the Hot-Cold Empathy Gap.

are more valuable in colder temperatures), Conlin et al.'s dual hypotheses are that the likelihood of returning the item is (1) declining in the order-date's temperature (order-date hypothesis), and (2) increasing in the return-date's temperature (return-date hypothesis). In other words, a lower order-date temperature or a higher return-date temperature is associated with a higher probability of returning the item. As a consequence of these temperature differentials, an individual's likelihood of returning the item after having received it should increase, which would demonstrate a Projection Bias on both the day the item was ordered and the day it was received.

To test these hypotheses, the authors obtained sales data from a large outdoor-apparel company. The company provided usable, detailed information on over 2 million orders of weather-related items, including the zip code of the buyer, the order date, and whether the item was ultimately returned. Specifically, the company provided information about each item ordered, its order date, the date the item was shipped, whether the item was returned, and, if so, the date on which the company restocked the item. In addition, the company provided information on the five-digit zip code associated with the billing address, whether the shipping address was the same as the billing address, the price of the item purchased, whether the order was placed over the Internet, by phone, or through the mail, and whether the buyer used a credit card to purchase the item. The company also provided information that enabled the authors to construct a two-day window during which the buyer was most likely to have received the item. The authors then merged this data with daily weather information for each zip code in the US.

Conlin et al. ultimately find support for their order-date hypothesis—a decline in the order-date's temperature of 30°F is associated with an increase in the return rate of roughly 4%. However, they do not find support for the return-date hypothesis. Their specific econometric results are presented in the table below:

	With household fixed effects	Without household fixed effects
Order-date temperature	−0.00082**	−0.00039**
	(0.00027)	(0.00013)
Receiving-date temperature	0.00017	0.00002
	(0.00029)	(0.00015)
Average winter temperature, 1990–1994	0.00276**	0.00067**
	(0.00090)	(0.00019)
Days between order and shipment	−0.00710**	−0.00373**
	(0.00222)	(0.00114)
Days between order and receipt	0.00387*	0.00175
	(0.00214)	(0.00109)
Ordered through Internet	−0.01597	−0.00858
	(0.01521)	(0.00648)
Ordered by female	0.03454**	0.01398**
	(0.00647)	(0.00205)
First-time buyer	−0.09726**	−0.00913**
	(0.00983)	(0.00314)
Number of prior purchases	0.00016	−0.00013**
	(0.00011)	(0.00002)
Percent of prior purchases returned	−0.45905**	−0.06196**
	(0.02438)	(0.00618)
Price of item	0.00106**	0.00070**
	(0.00026)	(0.00015)
Purchased with credit card	0.05714**	0.02638**
	(0.01583)	(0.00741)
Items in order	0.00551**	0.00250**
	(0.00121)	(0.00053)
Clothing-type fixed effects	YES	YES
Item fixed effects	YES	YES
Month-region fixed effects	YES	YES
Year-region fixed effects	YES	YES
Household fixed effects	YES	NO
Observations	162,580	162,580
R-squared	0.19	0.10

Notes: For each column, the dependent variable is whether an item is returned (= 1 if item returned, and 0 otherwise), and the table presents the coefficient estimates from a linear regression. Standard errors are in parentheses—the standard errors are robust to arbitrary heteroskedasticity and correlation within a household.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

(*Conlin et al. 2007*)

As the table indicates, the coefficient estimate for order-day temperature, with and without "household fixed effects," is negative and statistically significant.[54] In other words, a lower order-date temperature is indeed associated with a higher probability that the item will later be returned (i.e. evidence of Order-Date Projection Bias). To the contrary, the authors do not find strong support for the return-date hypothesis. While the coefficient estimate on receiving-date temperature is positive and larger when household fixed effects are controlled for, the coefficients are not statistically significant.

Poor *Homo sapiens*. I don't know about you, but I find returning items I've previously purchased to be a hassle; a hassle *Homo economicus* never experiences.

54. Household fixed effects control for the unexplained variation in the relationship between whether an item is returned and a given household in the sample.

STUDENT PROCRASTINATION

Caplan and Gilbert (2008) define two different types of student procrastination—"late-starting" and "back-loading." As their names suggest, late-starting procrastination occurs when a student gets started on a homework assignment closer to the assignment's due date, and back-loading procrastination occurs when a student (who may have started early on an assignment) waits until later (closer to its due date) to finish the assignment. Assessing grades on a series of homework assignments completed by students in an intermediate microeconomics course, the authors found that both late-starting and back-loading procrastination reduce the typical student's score on any given homework assignment.

Using data compiled by the course's web-based course-management tool, Caplan and Gilbert (2008) were able to define late-starting according to the difference (in days) between an assignment's grading deadline and when the student first accessed the assignment online to answer one of its questions.[55] They defined back-loading according to skewness in the distribution of a student's time differences (in minutes) between an assignment's grading deadline and when a student first accessed each of the assignment's questions. The greater the extent of positive skewness in a student's time distribution, the more the student back-loaded the assignment.

The authors found that for each day a student late-started an assignment, his score fell by slightly less than 3%. Back-loading resulted in a slightly less than 2% reduction in the assignment's grade per unit of skewness. The results controlled for whether a student attempted an assignment's practice problems beforehand, the student's grade point average (GPA), total credits enrolled for the semester, gender, total number of hours worked per week at a wage-paying job, number of children under the age of 18 years living in the household, and which third of the semester each assignment was given. Caplan and Gilbert (2008) conclude that procrastinators, both late-starting and back-loading, tend to perform worse on graded assignments than their non-dillydallying counterparts.

STOPPING PROCRASTINATION DEAD WITH DEADLINES

In an attempt to better understand the propensity of *Homo sapiens* to control their procrastination through self-imposed deadlines (as a Commitment Mechanism) and whether this type of "binding behavior" can improve task performance, Ariely and Wertenbroch (2002) involved approximately 100 executive-education students at the Massachusetts Institute of Technology (MIT) in a semester-long field experiment.[56] Fifty-one students were randomly assigned to a "free-choice" treatment group where they were free to set their own deadlines for three short papers throughout the semester, and 48 students were assigned to a "no-choice" control group where they were assigned fixed, evenly spaced deadlines for the three papers.[57] The authors found that, on average, the deadlines set by the

55. At the time of the study, the tool was owned by Aplia Inc. Aplia sold the technology to Cengage Learning in 2007.

56. Schelling (1989) provides another example of a commitment mechanism designed to mitigate substance abuse. In Denver, CO, a rehabilitation center treats wealthy cocaine addicts by having them write a self-incriminating letter which is made public if they fail a random urine analysis. In this way, the rehabilitation center is serving as a neutral enforcer of the mechanism.

57. Four external constraints were imposed on students in the free-choice group regarding the setting of their deadlines: (1) students had to hand in their papers no later than the final class of the semester, (2) students had to announce (to the instructor) their deadlines prior to the course's second lecture, (3) the deadlines were final and irrevocable, and (4) the deadlines were binding, such that each day of delay beyond a deadline would cause a 1% reduction in the paper's overall

free-choice students were roughly 33, 20, and 10 days before the end of the course for papers 1, 2, and 3, respectively.

Overall, less than 30% of the deadlines were set by free-choice students for the final week of class (calculated as roughly 45 out of a total of 153 papers (3 papers x 51 students in the group)). The majority of deadlines were set prior to the final lecture (27% of the free-choice students chose to submit all three papers on the final day of class). The results suggest that students are willing to self-impose deadlines to overcome procrastination even when the deadlines are costly (in terms of a grade penalty for missing a deadline).

Ariely and Wertenbroch (2002) also compared grades across the free-choice and no-choice groups to see if flexibility in setting deadlines enabled members of the former group to attain higher grades than members of the latter group. The authors found that, on average, grades were higher in the no-choice group, suggesting that the free-choice students suffered from self-control problems, and although they used the deadlines to help overcome these problems, they did not set the deadlines optimally. The greater flexibility afforded the free-choice students ultimately led to worse performance.

The moral of Ariely and Wertenbroch's story? There is a potential tradeoff for *Homo sapiens* between stopping procrastination dead in its tracks and one's overall performance of the task at hand.

TESTING THE SMALL-AREA HYPOTHESIS

A question loosely related to why *Homo sapiens* procrastinate is, what factors influence an individual's motivation to bring goals to completion? As Koo and Fishbach (2012) point out, previous research suggests that the closer people are to reaching a goal, the more resources they are willing to invest to reach it. For example, in the context of reward programs, consumers are more likely to make a purchase, and make it sooner, if they are only a few purchases away from receiving the reward. *Homo sapiens* in general prefer actions that appear more impactful and thus increase their perceived pace of progress. Drawing from this logic, Koo and Fishbach set out to test the Small-Area Hypothesis, which proposes that how people monitor their progress toward goal completion influences their motivation.[58] In particular, the authors distinguish between the framing of progress in terms of completed actions versus remaining actions to complete a goal. They hypothesize that when people start pursuing a goal, a focus on accumulated progress (e.g., 20% completed) is more motivating than a focus on remaining progress (e.g., 80% remaining). Then, toward the end of their pursuit, a focus on remaining progress (e.g., 20% remaining) is more motivating than a focus on completed progress (e.g., 80% completed). In other words, directing attention to small areas increases motivation because the marginal impact of each action toward goal achievement then appears relatively larger.

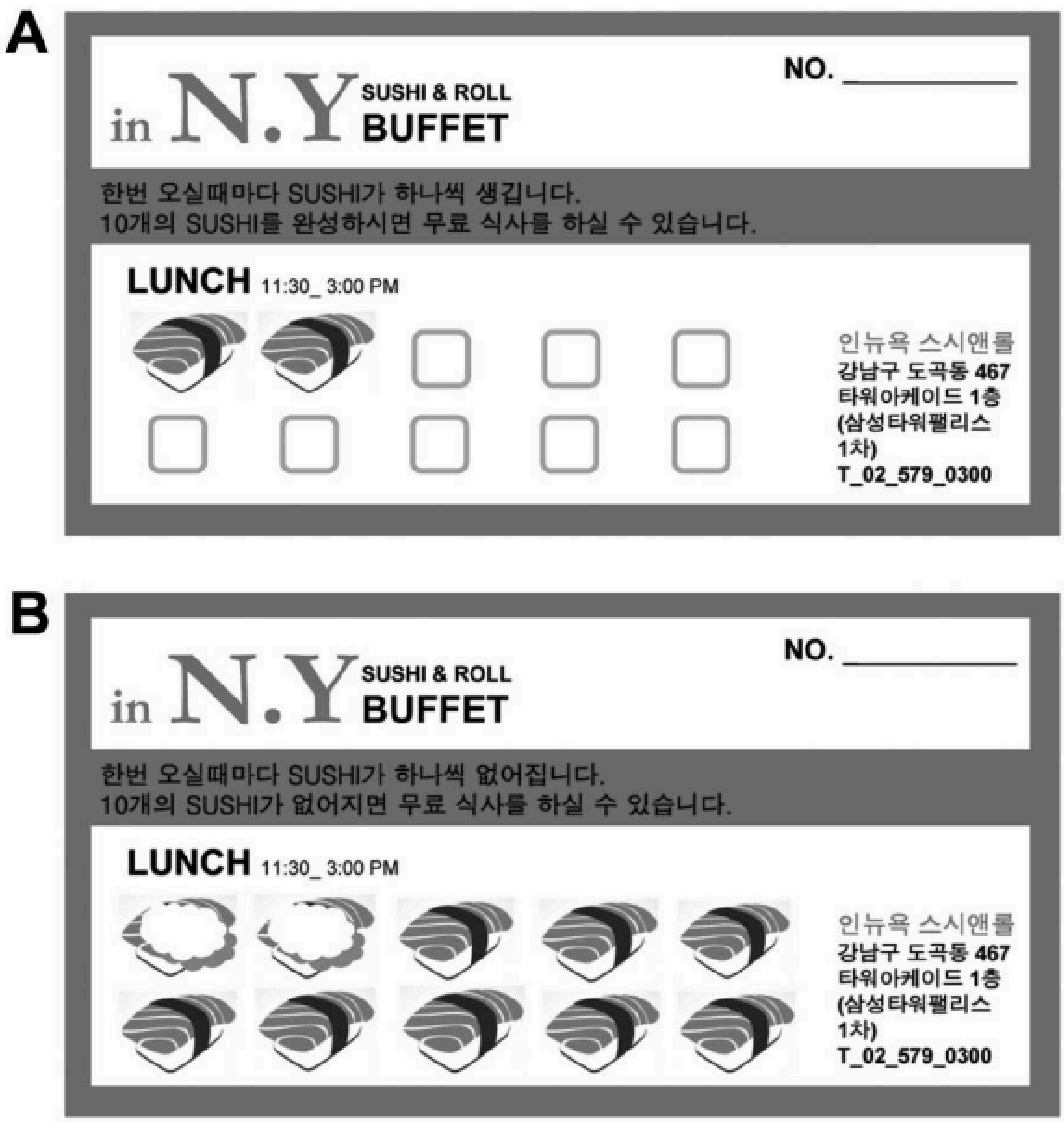
To test this hypothesis, Koo and Fishbach conducted a field experiment in a sushi restaurant that offered a buffet lunch menu located in a major metropolitan area in South Korea. For four months, they ran a reward program in the format of "buy 10 meals, get one free." The program manipulated customers' attention to accumulated versus remaining progress by providing them with a frequent buyer card on which they either received a stamp for each meal purchase (i.e., focus on accumulated progress) or had a slot removed for each meal purchase (i.e., focus on remaining actions needed). Over the four-month period, the researchers issued 907 reward cards corresponding to 907 participants,

grade. The authors argue that these constraints encouraged the free-choice students to submit all three of their papers on the last possible day of the semester.

58. See Martin et al. (2014) for a deep dive into the Small-Area Hypothesis.

though some customers may have redeemed one card and then received another, leading them to participate in the study more than once. The restaurant served a lunch sushi buffet for 20,000 won (US $18) per person. Koo and Fishbach offered the reward program solely for lunch buffet meals.

Participants in the field experiment received a frequent-buyer card similar to the ones displayed in the figure below. Participants were randomly assigned to one of two conditions. In the accumulated-progress condition, participants received a card to which a sushi-shaped stamp was added from left to right for each purchase (Card A in the figure). Thus, participants in this condition were provoked to direct their visual attention to the number of completed stamps. The text on the card indicated that customers would get a sushi stamp per every lunch meal and would be eligible for a free lunch meal once they received 10 stamps. In the remaining-progress condition, participants received a card with 10 printed sushi pictures already included (Card B in the figure). A punch was used to remove one picture from left to right for each purchase. In this condition, the participants were provoked to direct their visual attention to the number of remaining slots. The instructions on the card stated that a slot would be removed per every lunch meal purchased and that customers would be eligible for a free lunch meal after all 10 sushi pictures were removed.



(Koo and Fishbach 2012)

To assess a participant's initial progress, the authors recorded the number of purchases made at the time each reward card was issued. Because a single customer often paid for several lunches at the same time (covering the cost of her friends' meals), Koo and Fishbach were able to obtain natural variations in the level of progress customers made on their first visit. Overall, participants who attained a higher level of progress were more likely to use the card again (a result nevertheless driven by those who received a card with 10 printed sushi pictures already included). As the authors point out, this result reflects a Goal-Gradient Effect—those who made more purchases on the first visit were closer to the reward of the free meal and thus more likely to revisit the restaurant for a second time. Participants were more likely to revisit with the card highlighting remaining (vs. accumulated) purchases, possibly because attention to accumulated purchases encourages "resting on one's laurels."

Most importantly, these results indicate that high-progress participants (i.e., participants who purchased more meals initially) were more likely to revisit the restaurant if their card emphasized remaining purchases. Conversely, low-progress participants were more likely to revisit the restaurant with the card emphasizing accumulated purchases. As the authors state, this pattern supports the Small-Area Hypothesis—the higher the initial progress, the smaller the remaining-purchases area; and the lower the initial progress, the smaller the accumulated-purchases area.

Koo and Fishbach also measured how quickly participants revisited the restaurant by analyzing the number of days between the card's issuance and the participant's second visit to the restaurant. High-progress participants revisited the restaurant more quickly (i.e., within a shorter interval of days), a result driven mainly by those participants with cards emphasizing remaining purchases. Conversely, low-progress participants revisited the restaurant sooner if their card emphasized accumulated purchases. Hence, the Small-Area Hypothesis is again supported, this time regarding the amount of time elapsing between a participant's first and second visits to the restaurant.

The take-home message from this experiment is clear. If you are in the position of having to nudge yourself or someone else to accomplish a goal, you are more likely to attain that goal if you focus your persuasive messaging on the smaller of the two areas: the progress already made toward attaining the goal vs. the effort remaining to reach it. It is interesting to note that *Homo economicus*' behavior would not satisfy this hypothesis. It matters not which type of card is initially issued.

EXCESSIVE PLANNING

Setting personal goals and constructing respective plans to achieve them is a time-honored tradition of the *Homo sapiens* experience, so much so that a multitude of luminous thinkers have weighed in with memorable witticisms. A goal without a plan is just a wish (Antoine de Saint-Exupéry); By failing to prepare you are preparing to fail (Benjamin Franklin); If you don't know where you are going, you'll end up someplace else (Yogi Berra); and my three favorites, Plans are of little importance, but planning is essential (Winston Churchill); Always plan ahead. It wasn't raining when Noah built the ark (Richard Cushing); and, Good fortune is what happens when opportunity meets with planning (Thomas Edison).[59] One gets the distinct impression that the more planning the better.

Kirschenbaum et al. (1981) set out to test this impression in the context of a field experiment conducted with over 100 students at the University of Rochester during the spring semester of 1979. The students' goals were to leverage more structured and elaborate planning to enrich their studying time and ultimately improve their grades. The authors' goals were to test whether plans varying in specificity (i.e., detailed plans specifying a daily schedule ("daily plans") or looser plans outlining a monthly schedule ("monthly plans")), have differential impacts on the students' self-regulated study behaviors. The authors hypothesized that (1) students following daily plans would experience both greater process gains (i.e., with respect to developing more effective study habits) and performance gains (i.e., higher grades) than those following monthly plans; and (2) irrespective of whether they were following daily or monthly plans, students assigned to participate in an 11-session Study Improvement Program (SIP) would experience greater process gains than students in a control group that neither participated in the SIP nor devised plans for improving their studying effectiveness.

Kirschenbaum et al. recruited students with varying academic majors, half of whom were first-year undergraduates on academic probation or close to being on probation, who agreed to participate in the SIP. Students were first divided into "low-grade" and "high-grade" groups, the low-grade group

59. These quotes (and more) can be found on the website https://www.projectmanager.com/blog/planning-quotes.

consisting of students with cumulative grade point averages (GPAs) of 2.1 or less and the high-grade group with GPAs greater than 2.1. Roughly half of the low- and high-grade group members were then randomly assigned to respective control groups. The other halves were randomly assigned to one of six different treatment groups: daily plans-low grades, daily plans-high grades, monthly plans-low grades, monthly plans-high grades, no plans-low grades, and no plans-high grades.[60]

Students assigned to the daily and monthly plan groups used their course syllabi and textbooks to complete their plans (flow charts), which were then reviewed by other participants in the group and group leaders. The plans specified tasks to be accomplished, where and when they were to be worked on, the criterion of accomplishment, the self-administered reward to be earned, and where and when the reward was to be allocated. Daily-plan students completed highly specific plans in which they indicated study behaviors (e.g., activities, criteria, rewards) pertaining to each day (four days were planned on each flow chart). In contrast, monthly-plan students developed less-specific plans indicating larger chunks of activities that spanned one month's work on each flow chart, including a reward to be self-administered at the end of the month if the criterion level of accomplishment was achieved. Both groups were instructed to continue creating new plans when their flow charts became outdated (i.e., once a month for monthly plans and once every four days for daily plans). Daily-plan students also graphed their study time every day, while monthly-plan students graphed total study time each month. All participants self-monitored their study time on a daily basis.

Over the course of the ensuing semester, Kirschenbaum et al. found that monthly-plan students self-monitored more study hours and more "effective" (i.e., undistracted) study hours, and indicated less of a tendency to procrastinate, than did daily- and no-plan students. Monthly-plan students also maintained a relatively high rate of studying throughout the semester, whereas both daily- and no-plan students decelerated their study hours from the first to the second five-week period of the semester. While the authors found that within their cohorts both low-grade and high-grade students who participated in the SIP improved their GPAs relative to their respective control groups (i.e., low-grade SIP student GPAs improved relative to low-grade control group student GPAs and similarly for the high-grade SIP and control groups). No other statistically significant effects were found for student performance. High-grade students with daily and monthly plans did not perform better than high-grade students without plans, and similarly for low-grade students with and without plans.

Kirschenbaum et al. conclude that daily planning may have inhibited effective self-regulation by overburdening students with the task of planning for each day or by causing negative reactions in students who failed to meet their daily criteria for positive evaluation. In contrast, by writing out a proposed schedule for a longer period of time, monthly-plan students were able to increase their perception of control or choice. Either way, these results suggest that wise though it may be, students should be wary of the sage advice about planning far in advance.

KEEP YOUR OPTIONS OPEN?

Remember the old saying, Jack of all trades, master of none? It has historically (and pejoratively) distinguished someone who has developed relatively low-level competencies in a number of different occupations from one who has mastered the skills required of a single occupation. Each of us likely knows someone who has excelled in (i.e., mastered) their career at the expense of developing their skills in other facets of life, and someone who has dabbled in (i.e., jacked) different trades but never quite built a career in any one of them. Another example is someone who lacks hesitation when it

60. Students in the no plans-low grades and no plans-high grades treatment groups still participated in the SIP.

comes to making decisions—they tend to "jump right in"—versus someone who likes to keep their options open, and thus typically takes more time in making up their mind. *Homo sapiens* can indeed be fickle this way.

In an interesting set of experiments, Shin and Ariely (2004) investigate the extent to which *Homo sapiens'* penchant for keeping our options open leads to making inefficient choices. The authors ask whether the threat of future unavailability makes currently less-desirable options seem more appealing and whether this causes *Homo sapiens* to overinvest in these less-desirable options. In other words, do doors that threaten to close appear more attractive than doors that remain open? And if so, will individuals overinvest just to keep them open? As Shin and Ariely point out, we would expect *Homo economicus* to value an option (having the ability to make a choice) based solely upon the expected utility of the outcomes the option provides. To the contrary, we would expect *Homo sapiens* to be swayed by a preference for flexibility and by loss aversion, causing an option's subjective value to exceed its expected value.

Shin and Ariely's experiments all followed the same basic procedure. Participants interact with a computer game consisting of three different doors opening to three different rooms. One door is red, another blue, and the third green. By clicking with the mouse on one of the doors (i.e., by door-clicking), a participant opens the door and enters the room. Once in the room, the participant can begin clicking repeatedly in the room (i.e., room-clicking) to obtain points randomly drawn from a given distribution of points, or, after any number of room-clicks within one room, door-click another door (for no points on that click) to begin clicking in the new room for points randomly drawn from a different (albeit mean-preserving) distribution of points. The expected value and other moments of each distribution are unknown to the participants. Note that charging a participant a click for switching rooms creates a switching cost. The total number of clicks is prominently displayed on the computer screen, in terms of how many clicks have been used and how many remain until the end of the game.

Participants were initially given a "click budget" to use on door- and room-clicks at their discretion. Once participants used up all their clicks the game ended and they were paid the sum of their room-click payoffs. The main goal of the experiment was to measure the relationship between the actions of a participant and door availability (i.e., "option availability"), which varies on two levels: "constant availability" or "decreased availability." Under constant availability, all three rooms remain viable options throughout the experiment, irrespective of a participant's actions. Under decreased availability, door availability depends upon the actions of the participant. Each time the participant clicks on a door or within a room, the doors of the other two rooms decrease in size by 1/15th of their original width. A single door-click on a shrinking door re-sizes it to its original size and the process continues. If a door's size reaches zero it is eliminated for the remainder of the game. Poof!

Shin and Ariely's first experiment was designed to test the hypothesis that as doors shrink in the decreased-availability treatment, participants will invest in keeping their options open (i.e., spending a click to reverse a door's shrinkage). For this experiment, the expected value of a room-click (in any room) was 3 cents; however, the doors' point distributions were (1) Normal with a variance 2.25 cents and min/max 0/7 cents, (2) Normal with a variance 0.64 and min/max 1/5, and (3) Chi-Square with a variance 10 and min/max -2/10 for doors 1 – 3, respectively. Each participant was given a budget of 100 clicks. Over 150 students participated.

Shin and Ariely found that door switching was indeed significantly more likely to occur under decreased as opposed to constant availability. The authors also found that, overall, there was a decreased tendency among participants to switch rooms later in the game. However, more switching still occurred in the decreased-availability treatment. The authors point out that as click numbers increase, participants are gaining more experience, can better estimate the point distributions, and thus should have less of a need to explore other rooms (i.e., options). Further, the expected value of exploring other options is reduced with click number because the time horizon during which this constantly improving information can be put to use is being reduced.

It turns out that because of switching costs, picking a room and remaining there for the duration of the experiment is the optimal (rational) strategy in terms of earning the highest expected payout from the experiment (yes, this would be *Homo economicus*' strategy). Indeed, Shin and Ariely calculate that across both the constant- and decreased-availability treatments participants surrendered 11% of their payouts as a consequence of switching rooms (the average participant switched rooms 12 times during the course of the experiment).

Shin and Ariely's second experiment manipulated participants' knowledge about the point distributions for the various rooms, the hypothesis being that providing more information should substantially reduce the difference in switching between the constant- and decreased-availability treatments. On the other hand, if the information provided does not reduce this difference, then the difference is driven by a preference for flexibility and loss aversion rather than a paucity of information. The authors found evidence for the latter—*Homo sapiens* exhibit an inherent tendency to keep their options open—borne of a preference for flexibility in decision making or by loss aversion, even when doing so is costly.

ANTICIPATED VERSUS UNANTICIPATED INCOME

In a series of laboratory experiments, Arkes et al. (1994) tested whether *Homo sapiens* tend to spend (as opposed to save) unanticipated income (affectionately known as windfall gains) to a greater extent than anticipated gains (e.g., from owned assets). This is a non-issue for *Homo economicus*, who would never succumb to the temptation of spending more lavishly from unanticipated income. One experiment (Experiment 1) was designed to test whether students receiving unanticipated income would be more likely to risk that income on a simple lottery. The other experiment (Experiment 2) tested whether students receiving unanticipated income were more likely to spend the income on a consumer good. In each experiment, one randomly chosen group of students (the control group) arrived at the experiment anticipating some payment while the other group (the treatment group) was surprised by being given a payment upon arrival.

Students in both the treatment and control groups were informed between one and five days ahead of the experiment about the experiment's time and location. Students in the control group were also provided the following information:

> *Although it wasn't mentioned on the sign-up sheet for participating in the experiment, we want you to know that you will be paid for being in this experiment. We usually pay all our subjects $3.00 for participating. You will be paid when you get there. I thought you should know that. Also, I'd like to ask you not to mention to anyone that you're being paid. The reason for this is that not all psychology experiments pay the participants, so it's better if no one knows one way or the other.*

This additional information effectively distinguishes anticipated income (obtained by members of the control group) and unanticipated income (eventually obtained by members of the treatment group). Members of the treatment group unexpectedly received the $3 just prior to the start of the experiment. Unlike members of the control group, treatment-group members were not provided with any information prior to the experiment.

After receipt of the $3 the control- and treatment-group students participating in Experiment 1 were presented with the following gamble:

> *You can bet as much of your $3 on the roll of a pair of dice, from 25¢ to the entire $3. If you roll a number seven or greater, you win. If you roll a number less than seven, you lose. For example, if you bet $1 and you roll a number seven or greater, I will pay you $1. If you roll a number less than seven, you will pay me $1. You can roll the dice only once. Do you understand? How much do you want to bet?*

Students in the control group wagered an average of $1 on the gamble, while those in the treatment group wagered a statistically different average of $2.16. In other words, those students experiencing a windfall gain of unanticipated income wagered more than double the amount of income on the gamble than students whose gain in income was anticipated.

Control- and treatment-group students participating in Experiment 2 were distinguished in the same manner as in Experiment 1. However, in this case, the students were paid $5 rather than $3, and rather than subsequently facing a gamble, the students were sent to a basketball game. After the game, the amount of the $5 each student spent on concessions at the game was tallied.

Similar to the results from Experiment 1, students in the control group spent an average of just under 40¢ at the game while those in the treatment group spent a statistically different average of 90¢. This is yet another indication that *Homo sapiens* tend to be more spendthrift with windfall gains!

OPTIMISTIC OVERCONFIDENCE IN THE STOCK MARKET

As Barber and Odean (2001) point out, theoretical models of investor behavior predict excessive trading in the stock market by overconfident investors. Likewise, psychological research demonstrates that in areas such as personal finance, men are more overconfident than women (c.f. Karabenick and Addy, 1979). Together, these two theories predict that men will trade more excessively than women, a prediction the authors test by partitioning investors according to gender. Using account data for over 35,000 households from a large discount brokerage, they analyze the common stock investments of men and women over the course of six years, from 1991 through 1997. The authors find that men trade 45% more of their portfolios than women and that trading reduces men's net returns by 2.65 percentage points per year (relative to what their net returns would have been had they not traded) as opposed to a 1.72 percentage-point reduction in net returns for women traders.

Barber and Odean (2001) find that, across all households in the sample, men's average monthly turnover of stocks in their portfolio (technically defined as $(a \cdot b)/c$ where a = shares sold, b = beginning-of-month price per share, and c = total beginning-of-month market value of the owner's portfolio) was roughly 2% greater than the average woman's. This average difference in turnover rate between men and women was larger than that for the subsample of married households and lower than that for single-headed households. Across all households, women traders earn net monthly returns (what Barber and Odean call own-benchmark monthly abnormal net returns) that are 0.143% lower than those earned by the portfolio they held at the beginning of the year, while men traders earn net monthly returns that are 0.221% lower than those earned by the portfolio they held at the beginning of the year. Both shortfalls are statistically significant at the 99% confidence level, as is their difference of 0.078%. As with the stock turnover rates, the differences in net monthly returns between men and women traders are lower among married households and higher among single-headed households.

Bottom line for *Homo sapiens*? Not only are men overconfident in their investing acumen relative to women, but they also suffer larger losses in their investment portfolios than they otherwise would had they not been so confident in their investing abilities.[61]

THE EQUITY PREMIUM PUZZLE

As pointed out by Benartzi and Thaler (1995), historically there has been an enormous discrepancy between the returns on equity (e.g., stocks) and fixed-income securities (e.g., treasury bills and bonds). Since 1926, the annual real return on stocks has been roughly 7% while, for treasury bills, the return has been less than 1%. In an early attempt to explain the extent of this "equity premium," Mehra and Prescott (1985) found that plausible levels of investor risk aversion were an unlikely culprit, which, in turn, signaled the need for an explanation grounded in a framework other than that offered by the rational choice model of *Homo economicus*. What makes it so puzzling is not so much that the premium exists. Rather, given the premium exists, why would investors ever choose to hold securities?

Benartzi and Thaler offer an explanation for this puzzle that is firmly rooted in Prospect Theory. Investors are by nature loss averse, and long-term investors evaluate their portfolios frequently. Together, these two traits of *Homo sapiens* investors lead to what the authors call myopic loss aversion. In conjunction with being loss averse, the more often he evaluates his portfolio (i.e., the shorter his evaluative horizon), the less attractive an investor finds a high mean, high-risk investment such as stocks. By being loss averse, the investor effectively overreacts to the downside risk (i.e., when his stocks fall in value) and therefore over-invests in treasury bills or bonds.

The authors demonstrate how myopic loss aversion among investors can explain the equity premium puzzle by answering the question, "If investors exhibit myopic loss aversion, how often would they have to evaluate their investment portfolios in order to explain the equity puzzle" (page 81)? To answer this question, Benartzi and Thaler draw samples from the historical (1926-1990) monthly returns on stocks, bonds, and treasury bills provided by the Center for Research in Security Prices (CRSP). Using this data, they can simulate prospective utility levels associated with portfolios holding purely stocks and purely bonds for evaluation periods starting at one month and increasing one month at a time. The authors find that the evaluation period at which stocks and bonds are equally attractive, i.e., where the common portfolio consisting of 50% stocks and 50% bonds is optimally held, occurs at roughly 13 months. In other words, the common portfolio is most likely to occur when the representative investor evaluates stock and bond returns roughly once per year.

The moral of this story is striking. Because stock returns have historically outperformed bond returns, there is no reason to believe that a pure stock portfolio (or at least a portfolio heavily weighted in stocks) should not continue to outperform a portfolio more heavily weighted in bonds. Therefore, if you are a *Homo sapiens* prone to suffer from myopic loss aversion, it is best to invest in a stock

61. Oskamp (1965) devised a field experiment to similarly investigate whether psychologists' confidence in their own clinical decisions is justified. In the experiment, a group of over 30 psychologists read background information about a published case study that they were previously unfamiliar with. After reading each section of the study, the subjects answered a set of questions involving their personal judgments about the case. Results strongly supported the existence of overconfidence. Accuracy did not increase significantly with increasing information, but self-confidence in their judgments increased steadily and significantly. Oskamp concludes that increases in self-confidence do not necessarily portend increasing predictive accuracy about a given case.

portfolio at the outset, and then avoid evaluating your returns for as many years as you can. Yes, this is a case where procrastination (in reviewing your stock portfolio) actually pays off!

If constraining yourself to this extent places too heavy a burden on your curiosity, then at the very least consider practicing what Galai and Sade (2006) and Karlsson et al. (2009) have coined the Ostrich Effect. This effect occurs when an investor evaluates his returns more often after a rise in the stock market (i.e., after receiving good news) than after receiving bad news about a fall in the market. By practicing the Ostrich Effect an investor helps to offset the inherent, negative impacts of myopic loss aversion.

ENDOWMENT EFFECTS AMONG EXPERIENCED VERSUS INEXPERIENCED TRADERS

In Chapter 6, a simple laboratory experiment was proposed to test for an Endowment Effect in a constructed market setting. We surmised that in a market characterized by a relatively strong endowment effect exhibited by the sellers, one would expect most sellers' WTA values to exceed buyers' WTP values, resulting in few sales ultimately being consummated. In other words, to the extent that *Homo sapiens* sellers betray an Endowment Effect, we would not expect them to behave like their dispassionate *Homo economicus* counterparts, who, by virtue of their immunity to such an effect, would likely consummate more sales in the market and thereby generate larger gains from trade.

Experimenting in a constructed market setting has its advantages, foremost among them the ability of the experimenter to mitigate potential confounding factors correlated with any given real-world context. However, as List (2004) demonstrates, when it comes to testing for an Endowment Effect among sellers of an everyday consumable good, a well-functioning, real-world marketplace provides an ideal setting for a field experiment. In such a marketplace (which in List's case is a sports card trader's show), List can distinguish experienced from inexperienced sellers and thus measure the divergence in the strength of the Endowment Effect among these two seller types. In his experiment, List's "everyday consumable goods" are candy bars and coffee mugs. He finds that inexperienced sellers exhibit a relatively strong Endowment Effect. Experienced sellers behave more like *Homo economicus*; they are capable of eschewing the endowment urge.

List designed two versions of his experiment—one in which the market mimics a typical private market where buyers and sellers interact in an uncoordinated setting (Experiment 1), the other in which a collective-choice mechanism guides the sellers' individual decisions toward a coordinated outcome (Experiment 2).

In Experiment 1, each subject was randomly assigned to one of four treatments. The treatments differ by type of endowment. Subjects in treatment E_{mug} are initially endowed with one coffee mug, subjects in treatment $E_{Candy\,bar}$ one candy bar, subjects in treatment E_{both} one mug and one candy bar, and subjects in treatment $E_{neither}$ neither a mug nor a candy bar. The coffee mug retailed for just under \$6 at the University of Arizona bookstore. The candy bar was an extra-large, fine Swiss chocolate bar that also sold for roughly \$6.00. Fundamental insights gained from the treatments come from subjects' choices when asked if they would like to trade their initial endowment (with the experimenter). A subject can either keep her initial endowment or trade for the other good (e.g., mug for candy bar if in treatment E_{mug}, candy bar for mug if in $E_{Candy\,bar}$, both the mug and candy bar for one or the other if in E_{both} (weird I know, but this helps serve as a control on the treatments E_{mug} and $E_{Candy\,bar}$), and if in treatment $E_{neither}$, the subject chooses either the mug or the candy bar from the experimenter (this treatment also serves as a control vis-a-vis E_{mug} and $E_{Candy\,bar}$).

List also ran the same four treatments in Experiment 2 using a collective choice mechanism.

The four collective-choice treatments use identical mugs and candy bars. In public good treatment $E_{Candy\,bar}$, for example, subjects are initially endowed with a candy bar and must vote on a proposition to fund Mr. Twister, a small metal box placed at the front of the room that dispenses mugs. If the group chooses to fund Mr. Twister via simple majority rule, all N subjects in the room are required to give their candy bars to the experimenter. Mr. Twister's handle is then cranked N times, which delivers N mugs. The other three treatments are the public-good analogues of treatments E_{mug}, E_{both}, and $E_{neither}$.

The author conducted some of the treatments with professional dealers and others with ordinary consumers, which allowed him to exploit the distinction between *Homo sapiens* who have more trading experience ("dealers") and those who have less trading experience ("non-dealers"). As List reminds us, under both individual and group choice, the rational model of *Homo economicus* and Prospect Theory (as applied to *Homo sapiens*) have disparate predictions about choice behavior across the various endowment points. For preferences overall to be consistent with the rational model, the proportion of subjects in treatment E_{mug} who trade their mugs for candy bars should be equal to one minus the proportion of subjects in treatment $E_{Candy\,bar}$ who trade their candy bars for mugs. For example, if 70% of the subjects in treatment E_{mug} decide to trade their mugs for candy bars, 30% of subjects in treatment $E_{Candy\,bar}$ should trade their candy bars for mugs. As a result, 70% of subjects in each treatment end up owning a candy bar.

List found that 81% of non-dealers in Experiment 1's $E_{Candy\,bar}$ treatment chose to keep their candy bars rather than trade for mugs. Similarly, 77% of non-dealers in Experiment 1's E_{mug} treatment chose to keep their mugs rather than trade for candy bars. Relative to the control treatments E_{both} and $E_{neither}$, the Prospect Theory prediction of an Endowment Effect among non-dealers holds. Similar results hold among non-dealers in Experiment 2 for the public good.[62]

The outcome was noticeably different for dealers in Experiment 1. In this case, the percentages of subjects holding onto their endowments in treatments $E_{Candy\,bar}$ and E_{mug} of 47% and 56% indicate a (statistically significant) absence of an Endowment Effect among dealers. Hence, List concludes that Prospect Theory demonstrates strong predictive power for inexperienced *Homo sapiens*. To the contrary, the traditional rational model of *Homo economicus* better predicts the behavior of *Homo sapiens* who already have considerable previous experience selling in a marketplace. Hence, when it comes to measuring the Endowment Effect, it seems that prior market experience is a key determining factor in predicting the choice behavior of *Homo sapiens*.[63], [64]

62. Experiment 2 was not conducted with dealers.

63. List (2006) finds similar results for dealers vs. non-dealers in the actual sports card market, where inter alia sports cards rather than mugs and candy bars are the tradable commodities.

64. Northcraft and Neale (1987) find weaker results for experienced vs. inexperienced subjects exhibiting an Anchoring Effect in a combination laboratory-field experiment, where amateurs (i.e., students) and experts (i.e., professional real estate agents) are tasked with valuing a property for sale in Tucson, Arizona. All subjects were provided with a brochure full of facts about the property, including a full set of visuals. The only attribute of the property differing across subjects was the property's reported list price. The property was listed at values slightly above or below $74,900, or slightly above or below $134,900, depending upon the brochure. The authors found that the higher list price anchored a significantly higher value assigned to the property by the subjects. Although the amateurs anchored their values to the reported list price more than the experts, the difference between the two groups was small. In follow-up questioning, the experts were less likely to admit to having anchored their values than the amateurs. Thus, Northcraft and Neale find that not only do experts anchor, but they also deny their susceptibility to the inevitable.

RELUCTANCE TO SELL IN THE STOCK MARKET

As Odean (1998) explains, the tendency among stock market investors to hold on to "loser stocks" for too long and sell "winner stocks" too soon is known as the Disposition Effect. This effect is a type of loss aversion whereby a *Homo sapiens* investor is averse to realizing a loss on the sale of stock whose market price has fallen below the investor's cost basis (i.e., the price at which the investor originally purchased the stock).

To test for the existence of a Disposition Effect among stock market investors, Odean obtained trading records for 10,000 accounts at a large discount brokerage house from 1987 through 1993. In analyzing these records, Odean finds that, overall, investors realized their gains from winner stocks more frequently than their losses from loser stocks. His analysis also indicates that a large number of investors engage in the tax-motivated selling of loser stocks, especially in December, in order to declare losses on their tax returns and reduce their income tax burden (i.e., he finds evidence of what one might call a Tax-Loss Declaration Effect).

Because there are competing explanations for why investors might sell their winners while retaining their losers, Odean's analysis simultaneously tests for these as well. For example, investors may simply believe that their current losers will in the future outperform their current winners. Thus, they may sell their winners to rebalance their investment portfolios (a Rebalancing Effect).[65] It could also be the case that investors refrain from selling losers due to higher transactions costs associated with trading at lower prices (Transaction Cost Effect). When the author controls for these two potential effects in the data he still finds evidence of a Disposition Effect, but not for Transaction Cost or Rebalancing Effects. Regarding the latter effect, Odean finds that the winners investors choose to sell continue in subsequent months to outperform the losers they keep. This result indicates that while investors may have an intent to rebalance their portfolios for improved performance, in general, they do not achieve this goal.

For insight into how Odean uses his data to test for the Disposition Effect, he provides the following example:

> "Suppose an investor has five stocks in his portfolio, A, B, C, D, and E. A and B are worth more than he paid for them; C, D, and E are worth less. Another investor has three stocks F, G, and H in her portfolio. F and G are worth more than she paid for them; H is worth less. On a particular day, the first investor sells shares of A and C. The next day the other investor sells shares of F. The sales of A and F are counted as realized gains. The sale of C is a realized loss. Since B and G could have been sold for a profit but weren't, they are counted as paper gains. D, E, and G are similarly counted as paper losses. So, [across these two investors over these two days, there were a total] of two realized gains, one realized a loss, two paper gains, and three paper losses counted. Realized gains, paper gains, realized losses, and paper losses are then summed for each account and across accounts" (p. 1782).

Odean then calculates the "proportion of gains realized" (PGR) and "proportion of losses realized" (PLR) according to the following two formulae:

$$PGR = \frac{Total\ Number\ of\ Realized\ Gains}{(Total\ Number\ of\ Realized\ Gains + Total\ Number\ of\ Paper\ Gains)}, \text{ and}$$

$$PLR = \frac{Total\ Number\ of\ Realized\ Losses}{(Total\ Number\ of\ Realized\ Losses + Total\ Number\ of\ Paper\ Losses)}.$$

From the example above, PGR = 0.5 and PLR = 0.25. As Odean points out, a Disposition Effect has unconditionally occurred if $PGR > PLR$ measured over the time horizon of analysis

65. This can also be thought of as investors holding a belief in "mean reversion" in terms of stock prices.

(1987-1993). Further, a tax-loss declaration effect has occurred if ($PLR - PGR$ in December) > ($PLR - PGR$ in January through November).

Odean finds that over the course of an entire year, $PGR > PLR$ and that the difference between the two is statistically significant. Thus, evidence supports the Disposition Effect among his sample of investors. Indeed, as Odean points out, the ratio of PGR to PLR is a little over 1.5, indicating that, all else equal, a stock that has increased in value is more than 50% more likely to be sold from day to day than a stock whose value has decreased. Further, the author finds that ($PLR - PGR$ in December) > ($PLR - PGR$ in January through November), again with correspondingly high statistical significance. This result is evidence supporting the Tax-Loss Declaration Effect.

RELUCTANCE TO SELL IN THE HOUSING MARKET

In the housing market, sellers incur a loss when they sell their house for less than they paid for it. Because housing markets typically consist of some sellers incurring losses and others incurring gains, the housing market, like the stock market, provides an opportune setting within which to test for the presence of a Disposition Effect among sellers. Genesove and Mayer (2001) performed this test using sales data from downtown Boston, MA between 1990 and 1997. Their data consisted of the price originally paid for an apartment (i.e., the purchase price) by an owner who later listed the apartment for sale, the price subsequently listed by the owner *cum* seller (i.e., the asking price), an estimate of the apartment's market value, and any outstanding loan at its sale (which enabled the authors to remove from their sample any sellers who would be reluctant to sell at a price below what they originally paid for the apartment because they had to repay the loan used for their original purchase).

The authors found that sellers do indeed exhibit a Disposition Effect. Sellers who are expected to make a loss on the sale of their apartment set a higher asking price, all else equal. Genesove and Mayer's (2001) main results are presented in Columns (1) and (2) in the table below. The two variables of most interest to us are LOSS and LTV. Variable LOSS is the difference between a seller's purchase price, on the one hand, and the apartment's estimated value in the quarter it is listed, or zero, whichever is larger. In other words, if a seller is facing a projected loss on the sale of her apartment, LOSS records the estimated extent of that loss for that seller. Otherwise, if a seller is not facing a projected loss on the sale of her apartment (i.e., the seller is facing a projected gain), then LOSS records a zero for that seller. Variable LTV is defined as the difference between a seller's loan-to-value ratio and 0.8, or zero, whichever is larger. In other words, if a seller's loan-to-value ratio is greater than 0.8, then LTV records the difference between that ratio and 0.8 for the seller. Otherwise, if a seller's loan-to-value ratio is less than 0.8, then LTV records a zero for that seller.

LOSS AVERSION AND LIST PRICES
DEPENDENT VARIABLE: LOG (ORIGINAL ASKING PRICE),
OLS equations, standard errors are in parentheses.

Variable	(1) All listings	(2) All listings	(3) All listings	(4) All listings	(5) All listings	(6) All listings
LOSS	0.35	0.25	0.63	0.53	0.35	0.24
	(0.06)	(0.06)	(0.04)	(0.04)	(0.06)	(0.06)
LOSS-squared			−0.26	−0.26		
			(0.04)	(0.04)		
LTV	0.06	0.05	0.03	0.03	0.06	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Estimated value in 1990	1.09	1.09	1.09	1.09	1.09	1.09
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Estimated price index at quarter of entry	0.86	0.80	0.91	0.85		
	(0.04)	(0.04)	(0.03)	(0.03)		
Residual from last sale price		0.11		0.11		0.11
		(0.02)		(0.02)		(0.02)
Months since last sale	−0.0002	−0.0003	−0.0002	−0.0003	−0.0002	−0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Dummy variables for quarter of entry	No	No	No	No	Yes	Yes
Constant	−0.77	−0.70	−0.84	−0.77	−0.88	−0.86
	(0.14)	(0.14)	(0.13)	(0.14)	(0.10)	(0.10)
R^2	0.85	0.86	0.86	0.86	0.86	0.86
Number of observations	5792	5792	5792	5792	5792	5792

(Genesove and Mayer 2001)

In Column (1) of the table, the statistically significant coefficient estimate for LOSS of 0.35 indicates that a 10% increase in a prospective loss leads a seller to set the asking price 3.5% higher, all else equal. The corresponding coefficient estimate for LOSS in Column (2) of 0.25 indicates that a 10% increase in a prospective loss leads a seller to set the asking price only 2.5% higher. Genesove and Mayer (2001) interpret the estimate from Column (1) as an upper-bound and that from Column (2) as a lower-bound on the true relationship between a prospective loss and the seller's asking price. Similarly, the statistically significant coefficient estimate for LTV of 0.06 in Column (1) indicates that a 10% increase in the loan-to-value ratio for those sellers with ratios already above 0.8 leads these more highly indebted sellers to set an (upper-bound) asking price just over 6% higher, all else equal. Column (2) indicates that the lower-bound on the relationship between LTV and a seller's asking price is 0.5%. Together, these results corresponding to LOSS and LTV indicate that sellers facing a loss on the sale of their apartment—both concerning their original purchase price and their loan-to-value ratio—exhibit a Disposition Effect by setting their asking prices above those set by sellers who do not face a loss.

So as not to shield buyers from their share of "*Homo sapiensism*" in the housing market—in particular *Homo sapiens*' affinity for reference dependence—Simonsohn and Loewenstein (2006) investigated the US housing market between 1983 and 1993 to discern whether the average monthly rent or house price in the location where households moved from (i.e., their old locations) affected the monthly rent or house price they paid in the location they moved to (i.e., their new locations). In other words, do households that paid more for their housing in their old locations on average pay more for housing in their new locations. The authors find that the higher the rent or price paid in the old location, the higher the rent or price paid in the new location. Further, when households move for a second time within their new location, this positive relationship between prices paid in the old and new locations disappears. Simonsohn and Loewenstein conclude that households readjust their reference points after having lived in an area for some time. Therefore, the moral of Simonsohn and Loewenstein's story is that even though *Homo sapiens* exhibit reference dependence in the housing market, at least their reference points are flexible.

DEAL OR NO DEAL?

When faced with an uncertain situation, do *Homo sapiens* set reference points based upon prior expectations, similar to how we set reference points in certain situations based upon prior experience? Post et al. (2008) set out to answer this question by assessing risky decisions made by 150 contestants from the Netherlands, Germany, and the US in the high-stakes game show Deal or No Deal.[66] In the game, contestants choose among 26 briefcases, each containing some uncertain amount of money, ranging from €0.01 to €5 million (in the Dutch edition of the game). Each contestant selects one of the briefcases and thereafter owns its unknown contents. Next, he picks six of the remaining 25 briefcases to open. Each of the opened briefcases reveals a prize that is not in the contestant's initially chosen briefcase. The contestant is then presented with a "bank offer" from the game's host, which is the opportunity to walk away with a sure amount of money based loosely upon the average amount contained in the remaining unopened briefcases (Deal?), or to choose another five briefcases to open, followed by another bank offer (No Deal?). The game continues in this fashion until the contestant either accepts a bank offer or rejects them all and walks away with whatever amount of money is in the initially chosen briefcase.

Post et al. (2008) find that the typical contestant's choices can be explained in large part by previous outcomes experienced during the game (e.g., the amounts of money in the opened briefcases and associated bank offers). Aversion to risk diminishes as prior expectations are either shattered by unfavorable outcomes (i.e., the opening of high-value briefcases) or surpassed by favorable outcomes (the opening of low-value briefcases)—known as Break-Even and House-Money Effects, respectively. This process of reference-point adjustment made by the contestant represents what the authors call path dependence, a form of dependence to which *Homo economicus* would never succumb.

Post et al.'s (2008) basic results for the Deal or No Deal contestants are presented in the table below. A contestant is labeled a Loser if his average remaining prize in the unopened briefcases (after having eliminated the lowest remaining prize) is among the worst one-third across all contestants in the same

66. To control for the potential cross-country confounding effects of culture, wealth, and contestant selection procedure (not to mention stake size and contestant behavior), the authors also conducted laboratory experiments with their students (a more homogeneous population).

round of the game. A contestant is a Winner if his average remaining prize is among the highest one-third, and Neutral if neither a Loser nor a Winner. The column titled %BO lists the bank offer as a percentage of the money amounts in the remaining unopened briefcases per round, No. indicates the number of contestants who take the bank offer (i.e., take the deal) per round, and %D indicates the percentage of contestants who take the bank offer per round.

	Loser			Neutral			Winner		
Round	% BO	No.	% D	% BO	No.	% D	% BO	No.	% D
A. Netherlands (N = 51)									
1	6	17	0	6	17	0	6	17	0
2	15	17	0	12	17	0	15	17	0
3	40	17	12	29	17	41	31	17	6
4	69	14	14	58	13	46	54	14	21
5	82	10	10	71	10	20	78	10	40
6	94	8	50	85	7	43	86	8	63
7	99	4	25	97	3	67	99	4	75
8	105	1	0	91	3	67	100	1	100
9	120	1	0	—	0	—	91	1	100
2–9		72	14		70	31		72	25
B. Germany (N = 47)									
1	7	9	0	7	9	0	8	9	0
2	16	16	0	13	15	0	14	16	0
3	35	16	0	33	15	0	33	16	0
4	46	16	0	44	15	0	47	16	0
5	65	16	0	54	15	13	57	16	0
6	83	15	0	67	15	20	66	15	27
7	107	13	0	80	12	25	76	13	15
8	117	11	0	89	11	55	86	11	36
9	107	8	38	106	7	57	98	8	50
2–9		111	3		105	17		111	13
C. United States (N = 53)									
1	9	18	0	10	17	0	13	18	0
2	19	18	0	19	17	0	25	18	0
3	41	18	0	29	17	0	39	18	0
4	57	18	0	42	17	0	51	18	0
5	69	18	0	55	17	6	62	18	0
6	78	18	11	68	16	31	73	18	11
7	92	15	27	87	13	23	84	15	53
8	94	9	22	95	10	70	87	9	56
9	92	4	50	101	6	67	99	4	50
2–9		118	8		113	18		118	14

(Post et al. 2008)

Focusing on the US sample of contestants (the results for the Netherlands and Germany samples are similar), we see that (1) %BO generally increases for each type of contestant (Loser, Winner, or Neutral) as the game progresses (i.e., the number of rounds increases), and (2) generally lower percentages of both Losers and Winners take the deal as compared to Neutrals as the game progresses. Overall, across rounds 2–9, 18% of all Deal or No Deal choices in the Neutral group are Deal, while only 8% and 14% of choices were Deal in the Loser and Winner groups, respectively. Post et al. (2008) interpret these results as evidence that risk aversion diminishes for both Losers and Winners, particularly for Losers, who have been unlucky in selecting which briefcases to open. Thus, prior outcomes are indeed an important reference points for risky choices.

HEALTH CLUB MEMBERSHIP

What do we do when *Homo sapiens* are naive about their time-inconsistent preferences? Using attendance data for close to 8,000 health club members in New England from 1997–2000, DellaVigna and Malmendier (2006) were able to test whether members prone to making time-inconsistent choices choose a membership plan that helps them overcome this tendency most efficiently.[67] In their sample, gym-goers have four different membership plans to choose from: (1) pay a $12 fee per visit, (2) pay $100 for a 10-visit pass, (3) sign an (automatically renewed, cancelable) monthly contract for unlimited visits at a standard fee of $85 per month, or (4) sign an annual contract (requiring in-person or in-writing renewal at the end of the contract) for unlimited visits at $850 per year. The authors find that, on average, members forego $600 in savings over the course of their memberships, indicating that they choose suboptimal membership plans given their attendance frequencies. DellaVigna and Malmendier attribute this suboptimality to optimistic overconfidence on the part of club members in terms of their future self-control or efficiency in attending the club. Sound familiar?

In particular, the authors find that members who choose a monthly membership pay on average 70% more than they would under the pay-as-you-go, fee-per-visit contract for the same number of visits. Eighty percent of these monthly members would have been better off had they paid per visit for the same number of visits. In addition, members who choose a monthly contract are 17% more likely to stay enrolled beyond one year than users committing upfront to an annual membership. Monthly members, therefore, end up paying higher fees for the option to cancel each month. Further, low-attendance members delay canceling their monthly contracts despite the small transaction costs of doing so.

Because of its automatic-renewal provision, the monthly contract's default position is "opt-out," meaning if a member decides to terminate the contract, she must opt out of it. By contrast, because of its non-automatic renewal provision, the annual contract's default position is "opt-in," whereby a member must opt into the contract on a yearly basis. In this way, the monthly contract is well suited for members who suffer from time inconsistency associated with procrastinating in joining the club, and then remembering to renew thereafter, while the annual membership better suits those members who have difficulty in motivating themselves to regularly attend the club for their workouts. For those members who end up attending less than they originally imagined they would, paying the per-visit fee is the best option. For those who follow through with attending often, the annual-fee membership seems to make the most sense. And for those who at the outset are unsure of how often they will attend, the monthly-fee membership seems best. DellaVigna and Malmendier (2006) investigate whether members choose the best membership plan for themselves at the outset and if not, then whether they learn and adjust to overcome their time-inconsistency problem.

The authors find that for monthly members there is not a month where their average price was beneath the standard $12 fee-per-visit or the $10 cost-per-visit associated with the 10-visit pass. On average, the price paid by the monthly members was above $17 per visit. Likewise, the average price paid by the annual members was above $15 per visit. Thus, on average, monthly and annual members are overconfident about their attendance at the club. They are not choosing their membership plans optimally.

Regarding the question of whether annual and monthly members learn and adjust to overcome their time-inconsistency problem, DellaVigna and Malmendier find that after the first year on an annual contract, the average annual member increases his monthly attendance to a point where the

67. It is common in the literature on time-inconsistency to distinguish between those Homo sapiens who are clever enough to account for (and thus overcome) their time-inconsistent tendencies, and those who are not. The former types are known as "sophisticates" and the latter as "naifs."

corresponding average price per visit falls from over $15 to approximately $11.30—lower than the $12 fee-per-visit, but still higher than the $10 cost-per-visit associated with the 10-visit pass. After the first six months on a monthly contract, the average monthly member decreases (yes, decreases) his monthly attendance to a point where the corresponding average price per visit rises to roughly $20 per visit. It appears that members with an annual membership adjust their attendance to an extent that mitigates the inefficiency of their choice but does not eliminate it. Members with monthly memberships, on the other hand, exacerbate the inefficiency of their choice. Alternatively stated, annual members learn to mitigate their time-inconsistency problem, while monthly members exacerbate theirs.[68]

LESSONS FROM AN 'ALL-YOU-CAN-EAT' EXPERIMENT

In Chapter 2, we were introduced to the notion of flat-rate pricing. Later, in Chapter 6, we encountered the Sunk Cost Fallacy. Conventional wisdom suggests that because All-You-Can-Eat (AYCE) restaurants charge a fixed price (i.e., flat rate) for a meal, the Sunk Cost Fallacy should be relatively easy to detect in *Homo sapiens* consumption behavior when they belly up to the buffet.

Because an AYCE customer faces zero marginal cost associated with additional amounts of food consumed, the rational model of *Homo economicus* suggests that he should continue to eat until the marginal utility of consumption reaches zero. This is because the per-unit cost of consumption continually decreases with the amount consumed. Once the AYCE customer has paid the fixed price for the meal, his budget constraint on added consumption is effectively obviated. The only thing stopping him now is his epidemiological and neurological impulses. As Just and Wansink (2011) point out, in an AYCE setting price can influence whether one chooses to eat at the restaurant, but it should not affect the amount of food one consumes once he has chosen to eat there.

As we know, some *Homo sapiens* are driven to "get their money's worth" in various situations (recall the experiments in Chapter 6 involving the choice of whether to drive through snowstorms and rainstorms to get to a sporting event). In other words, *Homo sapiens* are susceptible to the Sunk Cost Fallacy. To the extent that the flat-rate pricing of AYCE restaurants triggers the Sunk Cost Fallacy in their customers, increasing the price of an AYCE buffet should increase the amount of food a customer ultimately eats. Just and Wansink test this hypothesis by designing an innovative field experiment that assigned customers to one of two prices at an AYCE pizza buffet restaurant.[69] The authors find that those assigned to the higher-price treatment consumed just under 40% more pizza than those assigned to the lower-price treatment. In other words, a higher flat rate did indeed increase the amount of food consumed. But is it a Sunk Cost Fallacy that drove these results, or perhaps an alternative effect?

68. Miravete (2003) conducted a similar study with telephone customers regarding their choice of a calling plan. Customers in Miravete's sample had a choice between a flat-rate fee of $18.70 per month or a flat-rate of $14.02 plus per-call charges. Miravete found a high percentage of customers either over- or underestimated the number of calls made monthly. Roughly 40% of customers were in the wrong plan in the month of October, which fell to 33% two months later. Thus, like the annual members in DellaVigna and Malmendier's sample of health club members, phone customers on average learned to mitigate their time-inconsistency problem but not eliminate it.

69. Just and Wansink test an alternative hypothesis that might also explain a positive correlation between the flat-rate price and the amount of food consumed. The hypothesis, known as Positive Hedonic Price Utility, suggests that a higher fixed price in and of itself induces an AYCE customer to take more pleasure in the taste of the food. One reason could be that price is interpreted by the customer as a signal of quality, leading her to believe that the pizza is of higher quality because she has paid more for it. The authors do not find evidence to support this hypothesis.

Permission to conduct the experiment was granted by the Pizza Garden, an AYCE restaurant located one mile south of Champaign, Illinois. The experiment was conducted during the restaurant's exclusive lunch buffet hours on a Tuesday, Wednesday, and Thursday in early April 2005. Members of the experiment's control group paid for the pizza buffet at the regular price of $5.98, while members of the treatment group were given coupons for 50% off this regular price. A total of 66 subjects participated in the experiment.

As Just and Wansink point out, customers choosing to eat at this restaurant would have already decided to eat the buffet at the regular price. In fact, no individuals included in either the treatment or control group failed to purchase the pizza buffet. While in the restaurant, pizza consumption was measured by three assistants who served as hostesses. The assistants were blind to the purpose of the study and had no knowledge of which patrons had been randomly assigned to the control or treatment groups. The assistants noted how many pieces of pizza each customer brought back from the buffet table, and how much was left uneaten after each customer completed their meal. Because the assistants were also responsible for busing tables, collecting the uneaten food was possible to do without raising suspicion. Uneaten pizza was weighed in a back room to more accurately assess what percentage of the pizza taken from the buffet table was actually eaten. After paying for their meals, the experiment's participants completed a short questionnaire concerning their demographics, how much they believed they ate, and their quality assessments of the pizza. The size of the group each participant was a part of while eating their meal was also noted, as group size can be a determinant of how much an individual eats at a restaurant.

The authors find that, on average, participants paying the full price ate roughly one slice more than participants paying half price, which nevertheless resulted in the full-price participants paying roughly $0.58 more per slice than the half-price participants. Full-price participants also left more uneaten pizza on their plates as food waste. Therefore, Just and Wansink find evidence that higher prices do indeed lead to greater pizza consumption. The Sunk Cost Fallacy seems to be in play at AYCE restaurants.

THE PERSISTENCE OF POLITICAL MISPERCEPTIONS

Political misperceptions have probably existed for as long as *Homo sapiens* have practiced politics. Although their frequency, intensity, and the extent to which they are disseminated among the general public via social media outlets are worthy of concern, misinformation campaigns (and the unsubstantiated conspiracy theories they spawn) have a long history in worldwide politics.

In a series of field experiments with self-identified ideological subgroups of adults, Nyhan and Reifler (2010) investigated the extent to which corrective information embedded in realistic news reports can succeed in reducing prominent misperceptions about contemporary political issues (according to the authors, misperceptions occur when people's beliefs about factual matters are not supported by clear evidence and expert opinion—a definition that includes beliefs about the world that are both false and unsubstantiated). In each of the experiments, the subgroups failed to update their beliefs when presented with factually corrective information that runs counter to their ideological predispositions. In several instances, the authors find that the corrections actually strengthen (yes, strengthen) rather than weaken misperceptions among those most strongly tied to their predispositions.

Nyhan and Reifler premise their experiments on previous research showing that many citizens base their policy preferences on false, misleading, or unsubstantiated information related to their political ideologies and what they believe to be true. For instance, after the US invasion of Iraq in 2003, the belief that Iraq had stockpiled weapons of mass destruction prior to the invasion was closely aligned with one's level of support for President Bush. As the authors point out, people are typically exposed to corrective information within objective news reporting, pitting two sides of an argument against each other. Nevertheless, we *Homo sapiens* are likely to resist or reject arguments and evidence contradicting our opinions.

Specifically, Nyhan and Reifler test three hypotheses about the extent to which corrective information overrides, or at least tempers, the effect of a subject's political ideology:

- **Hypothesis 1.** The effect of corrective information on misperceptions will be moderated by political ideologies.
- **Hypothesis 2**. Corrective information will fail to reduce misperceptions among the ideological subgroup that is likely to hold the misperception.
- **Hypothesis 3**. In some cases, the interaction between corrective information and political ideology will be so strong that misperceptions will actually increase for the ideological subgroup in question. Ouch. This is known as a Backfire Effect.

In the experiments, subjects read mock newspaper articles containing a statement from a political figure who reinforced a widespread misperception concerning the war in Iraq, tax cuts, or stem cell research (three popular issues at the time in American politics). Subjects were randomly assigned to read articles that either included or did not include corrective information immediately after a false or misleading statement.

The first experiment tested the effectiveness of corrective information embedded in a news report on beliefs that Iraq had stockpiled weapons of mass destruction (WMD) immediately before the US invasion. As Nyhan and Reifler point out, one possible explanation for why this misperception persisted was that journalists had failed to adequately fact-check Bush administration assertions that the US had found WMDs in Iraq. Another was people's fear of death in the wake of the September 11, 2001, terrorist attacks, known as "salience of mortality."

Subjects were instructed to read a mock news article attributed to the Associated Press that reported on a Bush campaign speech in Pennsylvania in October 2004. As Nyhan and Reifler describe it, the article describes Bush's remarks as a rousing, no-retreat defense of the Iraq War. The article included a quote from Bush: "There was a risk, a real risk, that Saddam Hussein would pass weapons or materials or information to terrorist networks after September 11th, that was a risk we could not afford to take (page 312)." A control group received only this information, while one treatment group also received corrective information based upon the Duelfer Report, which documented the lack of both Iraqi stockpiles of WMDs and an active production program immediately prior to the US invasion. Another treatment group received a mortality-salience question: "Please briefly describe the emotions that the thought of your own death arouses in you. Jot down, as specifically as you can, what you think will happen to you as you physically die and once you are physically dead."

All subjects were then asked to state whether they agreed with the summary statement: "Immediately before the US invasion, Iraq had an active weapons of mass destruction program, the ability to produce these weapons, and large stockpiles of these weapons, but Saddam Hussein was able to hide or destroy these weapons right before US forces arrived." Responses were measured on a five-point Likert scale ranging from 1 = "strongly disagree" to 5 = "strongly agree." To gauge a subject's ideological disposition, subjects self-identified according to a centered seven-point Likert scale, ranging from -3 = "strongly liberal" to 3 = "strongly conservative." An additive five-question scale measuring political knowledge used conventional factual questions.

As the results in the table below demonstrate, Nyhan and Reifler find that, as expected, more knowledgeable subjects were less likely to agree with the summary statement (the coefficient estimates for Political Knowledge of -1.133 and -1.081 are statistically significant), conservatives were more likely to agree with the statement (the coefficient estimates for Ideology of 0.347 and 0.199 are also statistically significant), but neither the corrective information (Correction) nor mortality salience question (Mortality Salience) had statistically significant effects on subjects' responses.[70]

	Model 1	Model 2
Correction	0.065 (0.191)	0.240 (0.196)
Ideology	0.347 (0.064)***	0.199 (0.082)***
Political knowledge	−1.133 (0.372)***	−1.081 (0.362)***
Mortality salience	0.280 (0.192)	0.271 (0.187)
Correction * ideology		0.359 (0.127)***
Constant	3.245*** (0.331)	3.156*** (0.323)
R^2	0.24	0.29
N	130	130

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (two-sided)

(Nyhan and Reifler 2010)

In Model 2, the introduction of the interaction term *Correction*ideology* tests whether the effect of the corrective information is moderated by subjects' political ideologies (Hypothesis 1), in particular whether the corrective information will be increasingly ineffective as subjects' political ideologies increase their susceptibility to misperception—in the case of justifying the Iraq War among politically more conservative subjects (Hypothesis 2)—or whether the corrective statement will backfire and that misperceptions will actually increase among politically conservative subjects (Hypothesis 3). Because the coefficient estimate on this interaction term is positive and statistically significant, it supports each of the three hypotheses. Looking more closely at their data, the authors find that the corrective information worked as expected. When exposed to the corrective information, very liberal subjects became more likely to disagree with the summary statement. No statistical effect was found for subjects describing themselves as liberal, somewhat left of center, or centrist. And conservatives became more likely to agree with the summary statement…kaboom, corrective information backfired!

The authors go on to test whether the Backfire Effect occurred because conservative participants distrusted the news source, Associated Press. They find that news source (New York Times vs. Fox News) has no impact on the result. Interestingly, when the context of the mock news article is changed from the 2004 Bush campaign speech to a 2005 Bush statement about Iraq, the Backfire Effect not

70. Recall that responses to the summary statement are measured on a five-point Likert scale ranging from 1 = "strongly disagree" to 5 = "strongly agree." Hence, all else equal, a negative(positive) coefficient estimate indicates less(more) agreement with the statement.

only disappears in general, but transforms into a Forward-Fire Effect—conservatives receiving the corrective statement are less likely to agree with the summary statement. This is evidence of a framing effect. Nevertheless, among those subjects who rated Iraq as the nation's "most important problem," the Backfire Effect persisted with the 2005 Bush statement about Iraq.

Nyhan and Reifler put forth two possible justifications for the Forward-Fire Effect. They point out that conservatives may have shifted their rationale for supporting the war in tandem with the Bush administration, which over time sought to distance itself from the WMD rationale for the war. By early 2006, national polls suggested a decline in Republican beliefs that Iraq had stockpiled WMDs before the US invasion. Another possible explanation is that conservatives generally placed less importance on the war by early 2006, and thus were less likely to counterargue the corrective information.

Lastly, the authors found a similar Backfire Effect when the issue at hand was misperceptions about the Bush tax cuts of 2001 and 2003. Regarding the issue of misperceptions regarding stem cell research in the early 2000s—a misperception held among liberals—Nyhan and Reifler find that corrective information worked (i.e., had a statistically significant negative effect on misperceptions) among subjects self-identifying as centrists and right-of-center, but failed to affect subjects identifying as liberal (i.e., left-of-center and beyond). Thus, in this experiment evidence is found in favor of Hypotheses 1 and 2, but not Hypothesis 3 concerning stem cell research. Thankfully, stem cell research is not an issue inspiring a Backfire Effect among liberals.

TEMPTATION AND SELF-CONTROL – THE CASE OF POTATO CHIPS

As Wertenbroch (1998) observes, we *Homo sapiens* often cave in to temptation (e.g., about the consumption of "vice goods" such as cigarettes, as opposed to "virtue goods" such as reduced-fat yogurt) against our own better judgment and self-interest. In dealing with temptation, realizing immediate utility from consumption conflicts with the longer-term utility associated with self-control. Self-rationing is a form of self-control that limits a consumer's stock of vice goods and thus the possibility of consuming them. Self-rationing imposes transactions costs on additional consumption and is perhaps an expression of attendant feelings of guilt. One way to test the extent of a consumer's self-control is to answer the question, Are consumers less likely to purchase larger quantities of a vice good than a virtue good in response to equal unit-price reductions? If the answer to this question is "yes," then consumers exhibit self-control in the face of temptation.

Using an experimental market approach, Wertenbroch tests whether vice consumers are less price sensitive than virtue consumers by examining consumers' demands for potato chips at two different quantity-discount price depths offered for a large-purchase quantity. The potato chips are framed as either 25% fat (relative vice good) or 75% fat-free (relative virtue good). Approximately 300 MBA students at Yale University participated in the experiment. The subjects were first shown a 6-oz. bag of an existing brand of potato chips as a reference package size. A questionnaire then offered them the opportunity to buy zero, one, or three 6-oz. bags of a new brand of potato chips at different prices per bag—$1 for a single bag and $2.80 for three bags (if the subject had been randomly assigned to the "shallow discount" treatment group), or $1.80 for three bags (if randomly assigned to the "deep discount" treatment group). A single bag represents the small size and three bags represent the large size. The new brand was described as having an innovative mix of ingredients and as currently being test marketed.

Subjects were informed that approximately one in 10 of those who completed their questionnaires would be randomly selected in a lottery to receive $10 in compensation for agreeing to participate in the experiment. To ensure that subjects would accurately reveal their demand for the chips, they were informed that the lottery winners would have to purchase (out of their $10 compensation payment) the respective amounts of potato chips they had chosen in the questionnaire at the given prices.

Wertenbroch finds that subjects who bought potato chips were more likely to prefer the large size when the chips were framed as 25% fat (again, the vice good) than as 75% fat-free (the virtue good). However, the probability of buying the large size under the virtue frame increased from 20% under the shallow discount to 65% under the deep discount (i.e., an increase of 225%). Under the vice frame, the corresponding increase was from 41% to 53%, an increase of merely 29%. Therefore, as the author points out, increasing the depth of the quantity discount was less effective in enticing vice consumers to increase their purchase quantities, suggesting that they self-imposed a rationing constraint as external price constraints were relaxed.[71]

DISHONESTY'S TEMPTATION

According to the National Retail Federation (NRF), customer and employee theft, fraud, and losses from other "retail shrink" in the US totaled just under $62 billion (or approximately 1.6% of total sales) in 2019, representing a 22% increase over the previous year (NRF, 2020). You read that correctly, $62 billion, with a "b." Interpreting retail shrink as the aggregation of consumers being dishonest with the businesses that supply our retail goods and employees being dishonest with the businesses that employ them, this $62 billion can be thought of as representing the monetary cost of dishonesty in the retail sector of the economy.[72]

Let's face it. Dishonesty is an inexorable part of the human experience, so inexorable that even *Homo economicus* can be expected to be dishonest in any given situation when the coldly calculated expected benefit of dishonesty outweighs its expected cost. To demonstrate dishonesty's pervasiveness among *Homo sapiens* (or, as Mazar et al. (2008) describe it, to measure the extent to which a little bit of dishonesty yields profit without spoiling one's positive self-view) Mazar et al. conducted a series of experiments enabling a comparison between the performance of participants in control conditions (in which the participants had no opportunity to be dishonest) with "cheating conditions" (in which participants had the latitude to cheat).

In the first experiment, the authors tested whether reminding participants of their standards for honesty would induce greater levels of honesty among them than among participants who were not preempted with such reminders. Over two hundred MIT and Yale students participated in the experiment, which consisted of multiple paper-and-pencil tasks appearing together in a booklet. To

71. Wertenbroch also conducted a market experiment where, after having categorized participants as either "hedonic" or "prudent" consumers based upon their answers to a Consumer Impulsiveness Scale, the participants stated how many packages of regular-fat (vice good) or reduced-fat (virtue good) Oreo chocolate chip cookies they wanted to purchase at each of 20 different package prices. The author hypothesized that if subjects use self-rationing as a self-control mechanism, then hedonic subjects (i.e., those with a high need for self-control) would be more likely than prudent subjects (i.e., those with a low need for self-control) to ration their purchase quantities of the regular-fat Oreos (i.e., that individual demand is less price sensitive for regular-fat Oreos than for reduced-fat Oreos among hedonic subjects but not among prudent subjects). Further, hedonic subjects do not generally prefer reduced-fat Oreos—that is, their virtue demand does not exceed their vice demand at any price. These hypotheses were confirmed.

72. Determining the full net cost of retail shrink is complicated. The long-run impact on businesses (e.g., the extent to which retail shrink shrinks businesses' future growth) would somehow need to be measured. Also, security costs incurred by businesses to prevent shoplifting and employee theft need to be accounted for. The (monetized) benefits obtained by shoplifters and employee thieves would then need to be subtracted from these costs.

begin, participants were asked to either write down the names of 10 books they had read in high school (no moral reminder) or the Ten Commandments (moral reminder) within a two-minute time limit.[73] Next, the participants were provided with a test sheet and an answer sheet. The test sheet consisted of 20 matrices, each based upon a set of 12 three-digit numbers. Participants had four minutes in which to find two numbers per matrix that added up to 10. An example matrix is depicted below.

1.71	1.80	2.91
4.67	4.83	3.03
5.80	5.08	4.28
6.38	5.19	4.55

The answer sheet was used by a participant to report her total number of correctly solved matrices. At the end of the session, two randomly selected participants earned $10 for each correctly solved matrix.

At the end of the four-minute matrix task, the experimenter verified each participant's answers. Participants in the experiment's two treatment (or recycle) groups indicated the total number of correctly solved matrices on their answer sheets, and then tore out the original test sheets from the booklet and placed them in their belongings (to recycle on their own later), thus providing these groups of participants with an opportunity to cheat.

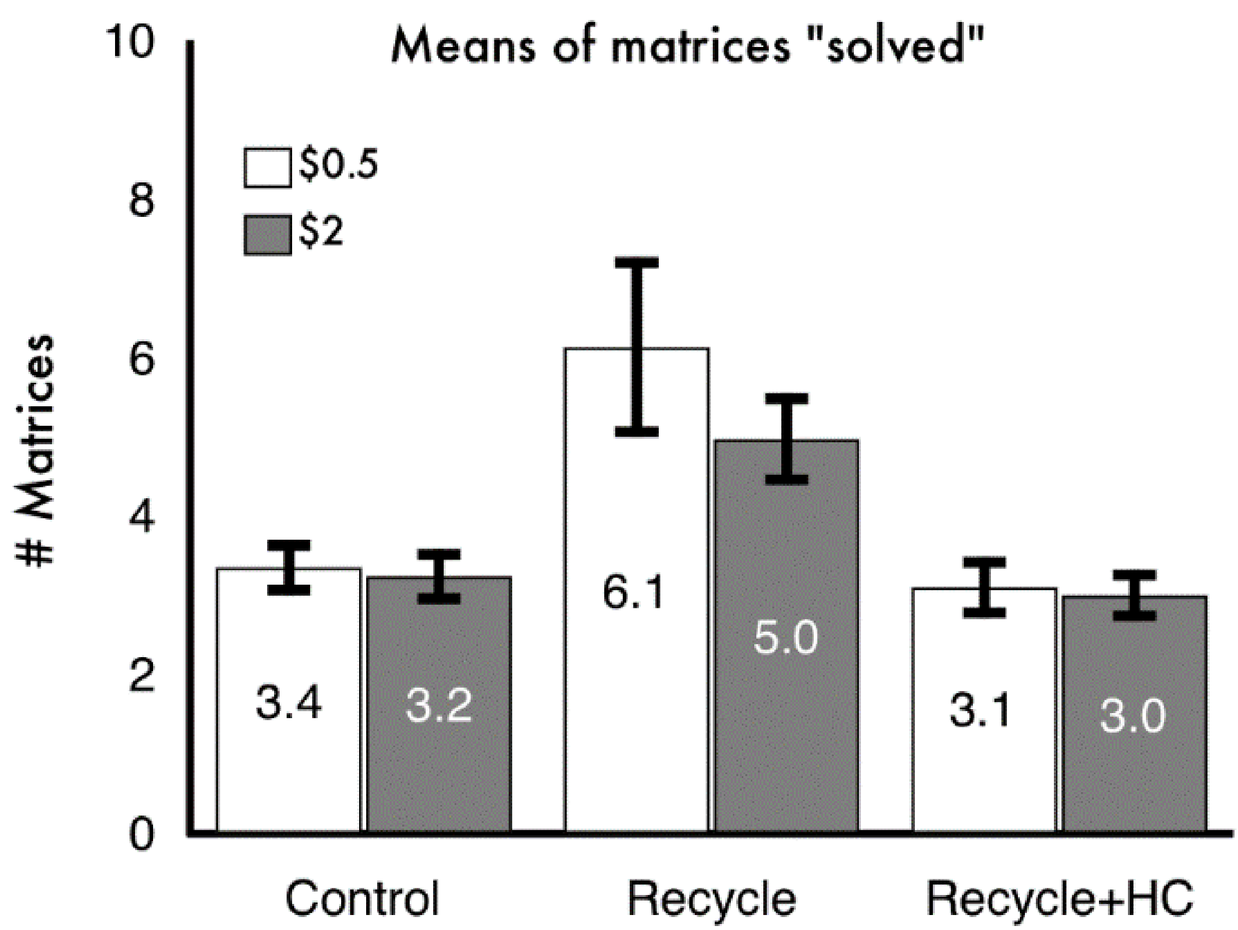
The results from this experiment were as anticipated. The type of reminder (10-books vs. Ten Commandments) did not affect the average participant's performance in the two control conditions—each group averaged just over three correctly solved matrices—suggesting that type of reminder influenced neither ability nor motivation. However, in the two treatment conditions, reminder type mattered. Following the 10-book recall task, participants self-reported an average of slightly more than four correctly solved matrices (which was significantly higher statistically than the control groups' three-matrix average), thus pointing to the likely presence of cheating among this group of participants. To the contrary, participants in the Ten Commandment recall task self-reported an average of slightly less than three correctly solved matrices (which was not significantly different statistically than the control groups' averages). Mazar et al. conclude that reminding participants of standards for morality eliminates cheating.

In the second experiment conducted with over 200 MIT and Yale students, the Ten Commandments recall task was replaced with an Honor Code treatment, the 10-book recall task was eliminated, and payments for correctly solved matrices (for randomly chosen participants) were

73. Students assigned to the Ten Commandments task recalled an average of slightly more than four of the commandments.

either 50 cents or $2 per matrix (to test for possible payment-level effects). In the two control groups, participants again handed both the test and answer sheets to the experimenter at the end of the matrix-solution task. The experimenter verified their answers and wrote down the number of correctly solved matrices on the answer sheet. In the two recycle treatments (one without any recall task, henceforth, "recycle" treatment, the other with the Honor Code recall task, henceforth, "recycle+HC") participants indicated the total number of correctly solved matrices on the answer sheet, folded the original test sheet, and then placed it in their belongings, similar to the first experiment. In the recycle+HC treatment, there was a statement located at the top of the matrices test sheet that read: "I understand that this short survey falls under MIT's [Yale's] honor system (page 637)." Participants printed and signed their names below the statement.

Results from this experiment are depicted in the figure below. Similar to the results for the first experiment, we see that while the recycle treatment resulted in a statistically significant increase in self-reported correctly solved matrices relative to the control groups and recycle+HC treatments (the whiskers do not overlap), the control groups and recycle+HC treatments did not result in statistically different scores. Interestingly, the different payment amounts (50 cents and $2) did not result in different scores within each respective group type—control, recycle, or recycle+HC.



(Mazar et al. 2008)

Lastly, in a third experiment, Mazar et al. tested whether the opportunity for dishonest behavior occurred in terms of money or in terms of an intermediary medium (tokens). The authors posited that introducing tokens would offer participants more latitude in interpreting their actions, hence making it easier for participants to justify cheating. Participants (450 MIT and Yale students) had five minutes each to complete the matrix task and were promised 50 cents for each correctly solved matrix. The same control and recycle treatment as in the second experiment were used, along with a "recycle+token" treatment where participants knew that each (self-reported) correctly solved matrix earned one token, which could be exchanged for 50 cents a few seconds later.

Similar to their previous findings, Mazar et al. found that the average participant in the recycle treatment reported having solved significantly more matrices than the average participant in the control group, suggesting the presence of dishonesty among the former group. Interestingly, introducing tokens as the medium of immediate exchange further increased the magnitude of dishonesty in the recycle+token treatment, such that it was significantly larger than that exhibited in the recycle treatment. This leads the authors to conclude that a medium, such as tokens, facilitates dishonesty, which helps explain the high levels of employee theft and fraud (e.g., stealing office supplies and merchandise, and putting inappropriate expenses on expense accounts) found in the US retail industry. As Ariely (2008) puts it, "what a difference there is in cheating for money versus cheating for something that is a step away from cash!" (page 299)[74]

BIGGER UNIVERSITIES, SMALLER SILOS

In Chapter 1, the concept of homophily and the Silo Effect were briefly explored. Left unanswered was the question of how the size and diversity of social choices (i.e., social ecology) affect the similarities between relationship partners (i.e., the extent of homophily existing between the two). As Bahns et al. (2010) point out, the initiation of an interpersonal or inter-organizational relationship is not only a dyadic process. The process is also influenced by the broader group of social contacts present in the local environment. It is the social ecology that shapes the kinds of communication and interactions that occur between the two partners, potentially hardening or softening the pretext for a silo effect.

Bahns et al. aver that, in general, when *Homo sapiens* have a choice, they tend to initiate and build relationships with partners who are similar to them. In their study, they compare the degree of similarity within dyads in a particular social ecology—a college campus—that varies in the size of the available pool of relationship choices. The authors compare dyads formed among students in public settings at a large state university to dyads formed in the same way in smaller colleges in the same state. Because students located at the larger universities can choose among a greater variety of fellow students, Bahns et al. hypothesize that these students will also be able to match their interests and activities more closely with partners than students located at smaller universities, which leads to a straightforward, albeit ironic, hypothesis. Greater human diversity within an environment leads to less personal diversity within dyads.

To test this hypothesis 110 students (55 dyads) were recruited from a large campus (the University of Kansas) and 158 students (79 dyads) from four small universities located in small eastern and central Kansas towns. To collect their data, experimenters visited each campus on a midweek day and located a public space where students were interacting with each other (e.g., the student union and a cafeteria). Naturally occurring dyads, which were randomly identified, were defined as any group of

74. See Erat and Gneezy (2012) and Fischbacher and Föllmi-Heusi (2013) for alternative perspectives on the emergence of dishonesty among Homo sapiens.

exactly two people who appear to be interacting in some way. The experimenter then administered a five-section questionnaire.

The first section gathered information about the students' socio-demographics and the nature of their relationship (e.g., how long they had known their dyad partner, how close they were, and how many hours per week they spent with the partner). The second section of the questionnaire asked about different social attitudes concerning abortion, religious observance, birth control, the importance of maintaining traditional husband-wife roles in a marriage, and capital punishment. The third section measured what the authors call "feeling thermometers" of attitudes toward/prejudices against five different social groups—Arabs, Black Americans, overweight people, gay men, and Jews. The fourth section measured health-related behaviors (e.g., tobacco use, alcohol use, and exercise). The fifth section measured the extent of agreement with what the authors call relational mobility statements, i.e., (1) "At this school, it is easy to meet new people," (2) "People at this school have few chances to get to know new people," (3) "It is common for me to see people on campus who are unfamiliar" (page 123), and psychological independence statements, i.e., (1) "If a person hurts someone close to me, I feel personally hurt as well," (2) "My close relationships are unimportant to my sense of what kind of person I am," and (3) "Even when I strongly disagree with group members, I avoid an argument" (page 124).

Bahns et al. find that dyads on the smaller campuses reported less relational mobility, implying greater perceived relationship opportunities on the large campus. However, no evidence is found that distinguishes the degree of psychological independence across the large and small universities. Similarly, there was no statistically significant difference across universities regarding length of relationships and amount of time spent together. However, dyads from the smaller universities rated their relationships as being closer than those from the large university.

Participants from small universities reported somewhat more conservative political beliefs, more prejudice toward Black people, more negative attitudes toward abortion, and more positive attitudes toward religion compared to participants from the large university. Participants from the large university exercised less, drank more alcohol, and smoked more tobacco than participants from the small universities.

The authors conclude that attitudes and behaviors are meaningful and important dimensions of social relationships in both social ecologies—students sort into dyads along these lines. Most importantly for this study, Bahns et al. find significantly greater degrees of similarity within dyads formed at the large university than at the small universities in terms of socio-demographics and social attitudes. In other words, greater diversity within the university environment leads to less personal diversity within dyads. As the authors state,

> "It cannot be surprising that size of opportunity leads to the ability to fine-tune the outcome. When opportunity abounds, people are free to pursue more narrow selection criteria, but when fewer choices are available, they must find satisfaction using broader criteria" (p. 127).

TIPPING POINTS

To the extent that a given population of *Homo economicus* is comprised of risk-averse vs. risk-neutral vs. risk-loving individuals (and those whose (time-consistent) discount rates are relatively small vs. relatively large), tipping points like those explored in Gladwell (2002) are possible in a variety of social settings, such as the spread of disease (e.g., HIV/AIDs and syphilis), crime (e.g., use of crack cocaine and Methamphetamines), fashion trends (e.g., wearing of Hush Puppy and Airwalk shoes), popular children's shows (e,g., *Sesame Street* and *Blues Clues*), and new technologies (e.g., fax machines and cellular phones) in epidemic (or geometric) proportions. Throw in *Homo sapiens*' predispositions for reference dependence, loss aversion, and the many effects and biases encountered in Chapters 1 and 2, and the proverbial stage is set for tripping over the myriad of tipping points lurking out there in the real world.

As Gladwell points out, three interconnected characteristics underpin the spread of epidemics: (1) contagiousness of the micro-organism, fad, idea, or behavior in question, (2) the dependence of big effects on relatively small causes, and (3) the suddenness of change (i.e., the presence of tipping points). Particularly when it comes to epidemics depending upon word-of-mouth, these characteristics can also be thought of, roughly and respectively, as (1) the "stickiness factor" of an initiating message, (2) the "law of the few" individuals with rare social gifts, and (3) the "power of context" (i.e., the recognition that *Homo sapiens* are quite sensitive to their environments).

Gladwell specifies the Law of the Few as consisting of a confluence of three types of individuals: connectors (gregarious and intensely social individuals who know lots of other people from different walks of life, more as acquaintances than friends); mavens (information specialists who not only gather information but also revel in the opportunity to spread the information to others); and salesmen (those with the ability to persuade others who are unconvinced about what they are hearing from mavens and connectors).

According to Gladwell, the power of context relates to the subtle, hidden, and often unspoken messages or cues that are transmitted in the run-up to a tipping point. This implies that an individual's behavior is, to varying degrees, a function of social context. In the case of crime, for example, subtle messages sent by broken windows in a community or graffiti and broken turnstiles in a subway station help create a social context suggesting that it is ok to commit crime here.

Messages, both spoken and unspoken, are sticky when they are memorable and ultimately compel the recipient of the message to take a targeted action. Stickiness in this sense relates to a message's effectiveness, similar to the previously encountered messages designed to reduce littering, environmental theft, drunk driving, and to promote energy conservation and better health care.

Gladwell provides several examples of epidemics that have adhered to the patterns identified above (e.g., contagiousness, dependence of big effects on relatively small causes, and the presence of tipping points). These examples include the massive and rapid-fire success of the late 1960s children's educational TV show *Sesame Street* (and the show it later spawned in the mid-1990s, *Blue's Clues*), the direct-marketing campaign of the Columbia Record Club in the 1970s, the surge of crime and its subsequent reversal on the NY City subway system in the mid-1980s, adoption by US farmers of new hybrid seeds in the late 1930s, teenage suicide in Micronesia in the mid-1960s through the 1980s, and even Paul Revere's midnight ride at the outset of the American Revolutionary War in the late 1700s.

Consider the epidemic of teenage smoking in the US. Gladwell ascribes primary billing in this epidemic's cause to salespeople (recall the law of the few), in particular extroverts, individuals who tend to be more rebellious and defiant and who make snap judgments and take more risks. These are people who are not perceived as being cool because they smoke, rather they smoke because they are cool. In effect, this epidemic's salespeople are also its tipping points, or "tipping people." Adolescents are naturally drawn to them.

According to Gladwell, the epidemic's stickiness factor occurs naturally. Because the smoking experience is so memorable and powerful for certain people, they cannot stop smoking—the habit sticks. Whether a teenager picks up the habit depends upon whether he comes in contact with a salesperson who effectively gives the teenager permission to engage in deviant acts. Of course, whether a teenager likes smoking cigarettes enough to keep using them depends upon a very different set of criteria. As Gladwell points out, nicotine is highly addictive but only in some people some of the time. Millions of Americans manage to smoke regularly and not get hooked. For these individuals, smoking is contagious but not sticky.

What to do? Gladwell suggests we might attack the epidemic from different, albeit self-enforcing, angles. One angle would be to prevent the salespeople from smoking in the first place. Another would be to convince all those who look to salespeople for permission to smoke that they should look elsewhere, to get their social cues from non-smoking adults. Further, as with other neurologically triggered addictions, zeroing in on combatting depression among teenagers would enable the exploitation of a critical vulnerability in the addiction process.

Regardless of which angle is emphasized in the effort to countervail the profuse tipping points in teenage smoking, it is never too late to consider new approaches in the campaign to control the epidemic. According to the American Lung Association (2020), every day, almost 2,500 children under 18 years of age try their first cigarette, and more than 400 of them will become new, regular, daily smokers. Of adolescents who have smoked at least 100 cigarettes in their lifetime, most of them report that they would like to quit but are unable to do so. If current tobacco use patterns persist, an estimated 5.6 million of today's youth under 18 will die prematurely from a smoking-related disease.

MAGICAL THINKING

It is (hopefully) safe to say that the majority of *Homo sapiens* distinguish both themselves and *Homo economicus* from superheroes (*Homo vir fortis*)—those with magical powers. However, the extent to which we *Homo sapiens* engage in magical thinking from time-to-time is perhaps less distinguishable. As Pronin et al. (2006) puts it,

> "Every so often, we may learn that someone we have wished ill actually has become ill, or that the sports team for which we are cheering has in fact gone and won the game. When such things happen, although we are far from causal, we may nonetheless experience a sense of authorship—a feeling that we caused the events we had imagined" (p. 218).

To investigate the prevalence of this type of magical thinking, the authors designed experiments to examine whether and when such experiences of everyday magic might arise.[75] They propose the formal hypothesis that belief in one's own magical powers can arise when we infer that we have personally caused events based upon perceiving a relation between our thoughts and subsequent events. One experiment tests whether college students might come to believe that they have caused another person pain through a voodoo curse when they have thoughts about the person consistent with such harm.

In this experiment, participants assumed the role of "witch doctor" in an ostensible voodoo enactment involving a confederate (a role-playing experimenter) as their "victim." The authors arranged for participants to encounter either a victim who was offensive (henceforth evil) or one who was neutral. Following this encounter, participants were instructed to stick pins in a voodoo doll representing the victim, in the victim's presence. The victim subsequently responded by reporting a slight headache, and participants were queried about their reactions to this reported symptom. The

75. We emphasize "everyday" here. As Pronin et al. point out, superstition and magical thinking are often observed in circumstances involving stressful and uncertain events. For example, college athletes show superstitious behaviors in sports competitions, and war-zone inhabitants similarly report magical beliefs about their personal safety.

experiment made possible the investigation of whether participants who harbor evil thoughts toward a victim are more likely than neutral-thinking participants to perceive that they caused the victim harm.

Slightly fewer than 40 residents of Cambridge, MA were randomly assigned to either a neutral-thoughts condition or an evil-thoughts condition. Each participant and confederate (a 22-year-old man) was greeted in a waiting area by the experimenter and escorted to the laboratory. The participant and confederate were seated at a table with a handmade twig-and-cloth voodoo doll lying on it. The experimenter explained that the experiment was designed to assess psychosomatic symptoms and physical health symptoms resulting from psychological factors, and that the study was investigating this question in the context of Haitian Voodoo. For background, the experimenter furnished both individuals with an abridged version of Cannon's (1942) *Voodoo Death*. This scientific account of how voodoo curses might impact physical health was included to bolster the plausibility of curse effects.

In the condition designed to induce evil thoughts, the confederate arrived at the experiment 10 minutes late, thus keeping the participant and experimenter waiting. When the experimenter politely commented that she was really glad he made it, he muttered with condescension: What's the big deal? He wore a T-shirt emblazoned with the phrase Stupid People Shouldn't Breed, and he chewed gum with his mouth open. When the experimenter informed the participant and confederate that they had been given an extra copy of their consent form to keep, the confederate crumpled up his copy and tossed it toward the garbage can. He missed, shrugged, and left it on the floor. Finally, while he and the participant read the *Voodoo Death* article, he slowly rotated his pen on the tabletop, making a noise just noticeable enough to be grating. Post-experiment interviews indicated that participants in the evil-thoughts condition were cognizant of many of these annoyances and found themselves disliking the confederate. Although the confederate was, by design, aware of these adjustments in his behavior, he was otherwise uninformed about the study's hypotheses.

After reading *Voodoo Death*, the participant and confederate were each asked to pick slips from a hat to determine who would be the witch doctor and who would be the victim. Both slips were labeled witch doctor, but the confederate pretended that his said victim. The confederate victim was then asked to write his name on a slip of paper to be affixed to the doll. Both victim and witch doctor then completed a page entitled Baseline Symptom Questionnaire that asked them to indicate whether they currently had any of 26 physical symptoms (e.g., runny nose, sore muscles, and/or headache). The confederate circled "No" for each symptom. To ensure that the participant knew the victim's purported health status, the experimenter verbally confirmed that the victim currently had no symptoms.

The experimenter then informed both individuals that reported cases of voodoo suggest that the witch doctor should have some time alone to direct attention toward the victim, and away from external distractions (before invoking the curse by pricking the voodoo doll), and she escorted the victim from the room. The participant was then asked to generate vivid and concrete thoughts about the victim but not to say them aloud. Afterward, the experimenter returned with the victim, who was again seated across from the participant. The participant was instructed to stick the five available pins into the doll in the locations of the five major weaknesses of the body: the head, heart, stomach, left side, and right side. Once the participant completed this task of piercing the doll, the victim was asked to complete a second symptom questionnaire (identical to the first). However, this time the victim invariably circled one symptom: a headache. He elaborated at the bottom of the page: I have a bit of a headache now. When asked to confirm this symptom, he averred with a slightly uncomfortable facial expression and the response "Yeah." The experimenter then stated that she would like to take some time with the victim to question him in detail about his symptoms but that she would first quickly ask the witch doctor some questions about his or her experiences in the experiment.

With the victim escorted from the room, the witch doctor was asked the following six questions:

1. "Did you feel like you caused the symptoms that the 'victim' reported, either directly or indirectly?"
2. "Do you feel that your practice of voodoo affected the victim's symptoms?"
3. "How much do you feel like you tried to harm the victim?"
4. "Do you feel that sticking the pins in the doll was a bad thing to do?"
5. "Did any negative thoughts about the victim pop into your head during the minute you had to yourself before the voodoo exercise?"
6. "Did you have any negative thoughts toward the victim before (or while) you did the pin pricks?" (p. 221)

Pronin et al.'s results were as expected. Witch doctors in the evil-thoughts condition were successfully induced to think ill of their victim; they reported significantly more negative thoughts about the victim than those in the neutral-thoughts condition. Most importantly, witch doctors in the evil-thoughts condition were more likely than those in the neutral-thinking condition to believe that they had caused the victim's headache. Witch doctors prompted to think evil thoughts reported feeling no more guilt than those prompted to think more neutrally about their victim. The authors conjecture that witch doctors saw the victim's headache as a just reward for his unpleasant behavior, and so they were not upset at having caused him pain. Ouch!

CONCLUDING REMARKS

As mentioned in this section's Introduction, the empirical studies and field experiments discussed here exemplify how behavioral economists have tested for the existence of the biases, effects, and fallacies underpinning *Homo sapiens'* choice behaviors, as well as the extent to which different implications of Prospect Theory (e.g., loss aversion, reference dependence, and the endowment effect), hyperbolic discounting, and mental discounting help explain these behaviors. Several of the case studies examined in this section also broach a host of contexts in which *Homo sapiens* exhibit socially degenerative behaviors (such as racial discrimination, criminal behavior, time-inconsistency, deadweight gift-giving, and procrastination), and empirically measure the extent of these behaviors in real-world situations.

Thankfully, the proverbial story does not stop there. Several of this section's case studies examine what Thaler and Sunstein (2009) call "nudges" to correct degenerative behaviors. For example, we explored ways in which the design of default options can be used to save lives, the extent to which basic-income and microfinance programs can help alleviate poverty, the extent to which simply raising awareness can help reduce racial discrimination, and how monetary reward/punishment schemes and information campaigns can (at least to some extent) mitigate social ills such as homelessness, food waste, and drunk driving, and promote improvements in such areas as energy conservation, public health, income tax compliance, and voter turnout. These types of nudges are the bridges between what behavioral economics teaches about *Homo sapiens*' quirks and consequent choice behaviors, on the one hand, and public policies that, with varying degrees of success, reorient these choice behaviors for the social good.

The number of organizations that have formed during the past decade to promote public policies incorporating insights from behavioral economics (i.e., to nudge) is impressive. For example, The Behavioral Insights (BI) Team began as a small agency of the United Kingdom's (UK's) government whose mission was to design innovative nudges to improve the workings of British society. Today the BI team is a global social purpose company whose projects span over 30 different countries. The BI Team's policy areas include finance, crime and justice, education, energy, the environment and sustainability, health and well-being, international development, taxation, and work and the economy.

GreeNudge is a Norwegian non-profit organization focusing on Norway's health-care system, specifically how to nudge consumers to choose healthier and more environmentally friendly foods in grocery stores, and, through an effort called Behaviourlab, apply behavioral science toward the realization of the United Nation's 17 Development Goals. In Peru, the nation's Ministry of Education has established MineduLAB, a laboratory designed to leverage lessons from behavioral economics to improve the country's educational policies. And the World Bank's Mind, Behavior, and Development Unit (eMBeD) is the spearhead of a worldwide network of scientists and practitioners working closely with governments and other partners to diagnose, design, and evaluate behaviorally informed interventions to eliminate poverty and increase social equity.

These organizations work to operationalize nudges similar to those we have studied in this section. Hopefully, the list of these types of organizations will grow over time reflecting the impact insights from behavioral economics can have on the collective will of the very species whose quirks and irrationalities serve as the basis of the economists' discoveries.

STUDY QUESTIONS

Note: Questions marked with a "†" are adopted from Just (2013), and those marked with a "‡" are adopted from Cartwright (2014).

1. Take a good look at the Airbnb website. If you have never visited this site, it is a marketplace for short-term rentals of apartments, homes, and even guest rooms in owner-occupied homes. To see the type of information displayed on the website, first, click on the "Anywhere" tab at the top of the screen. Then, in the "Where" box, type in the name of your hometown. Now type in hypothetical Check-In and Check-Out dates, and click on the "Search" button. You can now click on a few of the featured rentals and browse through the information provided about the rentals and hosts. Based upon what you can learn about these rentals, do you think there is enough information provided to empirically test for racial discrimination among people who book reservations through this site (recall the discussion in this chapter on peer-to-peer lending)? If "yes," explain how you might use the information to conduct your empirical test. If "no," then what additional information would you need to obtain from Airbnb in order to conduct your test of potential racial discrimination?

2. Recall the field experiments discussed in this chapter that were designed to test the effectiveness of monetary rewards in changing an individual's behavior (e.g., to improve

student and teacher performances and reduce substance abuse through "contingency management" programs). (a) Do you see anything that may be ethically wrong or socially degenerative with the use of monetary reward schemes like these? (b) Design a field experiment of your own to test the efficacy of using a monetary reward to either boost positive behaviors or reduce negative behaviors among a target population of people. In your design, be sure to clearly identify the target population, whether there are control and treatment groups (and what distinguishes these groups), and what outcome will support your hypothesis concerning whether the monetary reward was effective or not.

3. During the 2020 Democratic Party Primary season, Presidential candidate Andrew Yang proposed a basic-income program called the Freedom Dividend. His candidate website provides detailed information about the program. Read through the information provided on this website. What do you see as the pros and cons of a program like this? Explain your reasoning.

4. † You are considering buying gifts for two of your friends. Both friends enjoy playing video games. However, both have reduced their budgets for these types of purchases because of the temptation they cause. David is tempted to buy games when they are first released rather than waiting to purchase the games later once prices have fallen. To thwart this temptation, David has committed himself to spend no greater than $35 for any given game. Alternatively, Avita is tempted to play video games for long periods of time, causing her to neglect other important responsibilities in her life. To combat this temptation, Avita has committed herself to play video games only when she is visiting other people's homes. Would David be better off receiving a new game that costs $70 or a gift of $70 cash? How about Avita?

5. ‡ In this section, we were introduced to a study of how Minnesota worked to increase income tax compliance among its citizens. How do reference dependence and the overweighting of improbable events (recall the experiment discussed in Chapter 6) contribute to compliance?

6. Recently, while accessing the Wikipedia website, I was confronted with the following appeal that popped up and covered the bulk of the page: "To all our readers in the US, it might be awkward, but please don't scroll past this. This Saturday, for the 1st time recently, we humbly ask you to defend Wikipedia's independence. 98% of our readers don't give; they simply look the other way. If you are an exceptional reader who has already donated, we sincerely thank you. If you donate just $2.75, Wikipedia could keep thriving for years. Most people donate because Wikipedia is useful. If Wikipedia has given you $2.75 worth of knowledge this year, take a minute to secure its future with a gift to the Wikimedia Endowment. Show the volunteers who bring you reliable, neutral information that their work matters. Thank you." Suggested payment amounts that I could then choose were: $2.75, $5, $10, $20, $30, $50, $100, and "Other" amount.(a) Given what you have already learned about public goods (see Section 3, Chapter 8) and messaging/information campaigns in this chapter (e.g., reducing environmental theft, littering, drunk driving, and tax evasion, and promoting energy conservation in this chapter), comment on Wikipedia's fundraising strategy.(b) Can you

recommend any ways Wikipedia might improve upon its strategy? Explain why your recommendations could improve Wikipedia's fundraising performance.

7. ‡Suppose Donald never pays his taxes and resents having to transfer money to the government. His current wealth level is \$1.1 million. Donald interprets paying any amount of tax as a loss. His worst outcome would be that he chooses not to report any income (thus paying no tax) and ends up getting audited. His best outcome would be reporting no income and not getting audited. Suppose Donald (unwittingly or not) calculates his decision weight on being audited (w_a) as,$w_a = \frac{p^\delta}{\left(p^\delta+(1-p)^\delta\right)^{\frac{1}{\delta}}}$, and his decision weight on not being audited as $w_{na} = 1 - w_a$. Suppose further that p = 0.02 and δ = 0.69, resulting in w_a = 0.06 and w_{na} = 0.94. Note that because $w_a > p$ Donald is indeed overweighting the improbable event of being audited. Now suppose Donald's value function is given by, $v\left(x;r\right) = \begin{cases} \sqrt{x} + (x-r) & if \quad x \geq r \\ \sqrt{x} - 2.25\left(r-x\right) & if \quad r > x \end{cases}$ where x represents Donald's current wealth and r his reference point. At what reference point will Donald choose to not report any income?

8. To what extent should the results of Haney et al.'s (1973) simulated prison study serve to inform the current debate about prison reform? Explain.

9. What are some differences between New York City taxi drivers, on the one hand, and Uber and Lyft drivers on the other, that make the latter drivers less likely to exhibit a negative wage elasticity?

10. † Governments often require people to obtain insurance. For example, all drivers in the US are required to carry auto insurance to cover damages to others in the event of an accident. Homeowners are often required by banks to carry insurance on their homes. Why do these requirements exist? One characteristic of an overconfident person is that she is continually surprised when what she thought was unlikely or impossible comes to pass. What would happen in these cases if people were not required to insure? What problems might arise if governments also prepared for emergencies in a way that displays overconfidence? What mechanisms could prevent overconfidence in government action?

11. In this section, we learned about Banerjee et al.'s (2013) study of microfinancing in Hyderabad, India, in particular the extent to which this approach can potentially enhance the profitability of small businesses. Search the internet for a microfinance program implemented in another part of the world, and report on its approach to financing small businesses in its market area.

12. As we learned in this section, appropriately assigned default options can save lives and help employees save more money for retirement. Can you think of how airlines might harness this "Default-Option Effect" to help their customers reduce their environmental footprints when it comes to traveling by air?

13. † Some businesses thrive on *Homo sapiens*' proclivity for hyperbolic time discounting. For example, payday loan companies offer short-term loans with ultrahigh interest rates designed to be paid off the next time the person is paid. (a) Suppose you were considering opening a payday loan company. Given that hyperbolic discounters often fail to follow through on plans (e.g., they procrastinate and exhibit time inconsistency), how might you structure your loans to ensure earlier repayments from your customers? (b) Lotteries typically offer winners the option of receiving either an annual payment of a relatively small amount that adds up to the full prize amount over several years or a one-time payment at a steep discount. Describe how time inconsistency might affect a lottery winner's decision. How might a lottery winner view his decision over time?

14. Why might an owner of a health club choose not to offer a pay-per-visit option to her customers?

15. † Researchers have found that hungry people tend to have greater cravings for more indulgent foods (i.e., foods that are high in sugar, fat, and salt). Suppose you are creating a line of convenience foods— either snack foods or frozen foods. (a) Describe the circumstances under which most people decide to eat convenience foods. What state of mind are they likely to be in? Given this, what types of convenience foods are most likely to be eaten? (b) Describe your strategy for creating a line of convenience foods. Is there any way to create a successful line of healthy convenience foods?

16. Discuss the pros and cons of employers choosing to penalize poorly performing employees versus rewarding well-performing employees with bonuses.

17. What are some ways in which *Homo sapiens* can avoid contributing to the maligned Deadweight Loss of Gift Giving?

18. Consider monetary rewards and punishments, on the one hand, versus information campaigns on the other. Which "hand" do you think is more effective in nudging *Homo sapiens* toward better behavior/performance? Why?

19. In this section, we learned about a strategy to control food waste and a strategy to control procrastination. What do these two strategies have in common?

20. In this section, we learned about the emergence of Projection Bias in two different contexts—grocery shopping on an empty stomach and ordering winter clothing on a relatively cold day. Describe another context where the potential for Projection Bias is likely to emerge.

21. Recall Caplan and Gilbert's (2008) results for back-loading procrastinators. Compare the results with those Dr. Adam Grant attributes to "Originals" in this Ted talk.

22. One conclusion drawn from Benartzi and Thaler's (1995) study of the Equity Premium Puzzle was that the Ostrich Effect can actually serve as an antidote for *Homo sapiens* investors' myopic loss aversion. Can you think of a situation where the Ostrich Effect could instead hinder *Homo sapiens*?

23. Recall the Sunk Cost Fallacy depicted in Chapter 6 in the context of whether or not to brave a rainstorm to see a basketball game and described in this section in the context of how much to consume at an All-You-Can-Eat (AYCE) restaurant. In both of these instances, *Homo sapiens* who succumb to this fallacy are implicitly assumed to suffer its consequences—take an undue risk to attend a basketball game and overeat at the AYCE restaurant. Can you think of an instance where succumbing to the fallacy could instead result in favorable consequences? Explain.

24. Referring to Nyhan and Reifler's (2010) study of the persistence of political misperceptions, if you were to define a typical individual's value function over "accurate perception" (i.e., gains) and "misperception" (i.e., losses), what would it look like? What do you consider to be the traits of a "typical individual" in terms of how they value accurate perceptions and misperceptions?

25. Do you think the results from Wertenbroch's (1998) potato chip study would be replicated if a field experiment was conducted with casino patrons, where the vice good is higher-stakes gambling at the craps, blackjack, or poker tables, and the virtue good is low-stakes gambling at the slot machines? How would you even design a field experiment to be conducted in a casino to answer Wertenbroch's main research question: Are consumers less likely to purchase larger quantities of a vice good or a virtue good in response to equal unit-price reductions?

26. This section presented several examples of how well-intentioned incentives can sometimes lead to perverse outcomes—recall the perverse long-term impact of monetary incentives on student performance in the Chicago public school system, the unexpected response of tardy parents at an Israeli pre-school, and the perverse outcomes associated with health care report cards in the US and targets for emergency ambulance services in the UK. Can you think of another example of an incentive system gone bad? Describe how it went bad.

27. In the section on Tipping Points, it was pointed out that even *Homo economicus* can trip over tipping points. Throw in *Homo sapiens*' predispositions for reference dependence, loss aversion, and the many effects and biases encountered in Chapters 1 and 2, it was stated, and the proverbial stage is set for tripping over the myriad of tipping points lurking out there in the real world. Discuss how reference dependence and loss aversion can expedite the process of tripping over a tipping point. Discuss how the Endowment Effect and Status Quo Bias might work to delay the tripping process.

28. Suppose you own a small coffee shop in a busy metropolitan area. You decide to initiate a punch-card reward program to help build a loyal customer base. How should you train your

baristas to exploit the Small-Area Effect?

29. Do you see any connection between Magical Thinking and Confirmation Bias (that you learned about in Chapter 2)? Why or why not?

30. Do you see any connection between conceptual information's effect on one's consumption experience and Confirmation Bias (that you learned about in Chapter 2)? Why or why not?

31. Do you see any connection between the arguments given for not "keeping your options open" and the Burning Bridges game of Chapter 7? Explain.

32. Is Willingness to Accept Pain necessarily a better measure of pain tolerance/threshold than Willingness to Avoid Pain? Explain.

33. What is it about the use of tokens as direct payment, rather than money, that induces more dishonesty among *Homo sapiens*? Based upon what you have learned from Mazar et al.'s dishonesty experiments, how might a store owner go about reducing theft among his or her employees?

Media Attributions

- Figure 10 (Section 4) © Gwyn Bevan and Richard Hamblin; 2009 Royal Statistical Society is licensed under a All Rights Reserved license
- Figure 11 (Section 4) © Gwyn Bevan and Richard Hamblin; 2009 Royal Statistical Society is licensed under a All Rights Reserved license
- Figure 12 (Section 4) © Gwyn Bevan and Richard Hamblin; 2009 Royal Statistical Society is licensed under a All Rights Reserved license
- Figure 13 (Section 4) © Nava Ashraf, Dean Karlan and Wesley Yin; 2006 Oxford University Press is licensed under a All Rights Reserved license
- Figure 14 (Section 4) © Kela; Kansaneläkelaitos - The Social Insurance Institution of Finland is licensed under a All Rights Reserved license
- Figure 14a (Section 4) © Stephen Knack and Philip Keefer; 1997 Oxford University Press is licensed under a All Rights Reserved license
- Table 2a (Section 4) © Stephen Knack and Philip Keefer; 1997 Oxford University Press is licensed under a All Rights Reserved license
- Figure 15 (Section 4) © Erez Yoeli, Moshe Hoffman, David G. Rand and Martin A. Nowak; 2013 National Academy of Sciences is licensed under a All Rights Reserved license
- Figure 16 (Section 4) © Erez Yoeli, Moshe Hoffman, David G. Rand and Martin A. Nowak; 2013 National Academy of Sciences is licensed under a All Rights Reserved license
- Figure 17 (Section 4) © Erez Yoeli, Moshe Hoffman, David G. Rand and Martin A. Nowak; 2013 National Academy of Sciences is licensed under a All Rights Reserved license
- Figure 17a (Section 4) © James Heyman and Dan Ariely; 2004 Association for Psychological Science is licensed under a All Rights Reserved license
- Figure 18 (Section 4) © Leonard Lee, Shane Frederick and Dan Ariely; 2006 Association for Psychological Science is licensed under a All Rights Reserved license
- Figure 19 (Section 4) © Eric J. Johnson and Daniel Goldstein; 2003 American Association for the Advancement of Science is licensed under a All Rights Reserved license
- Figure 20 (Section 4) © Eric J. Johnson and Daniel Goldstein; 2003 American Association for the Advancement of Science is licensed under a All Rights Reserved license
- Figure 22 (Section 4) © Silpa Kaza, Lisa C. Zao, Perinaz Bhada-Tata, and Frank Van Woerden is licensed under a CC BY (Attribution) license
- Figure 23 (Section 4) © Brent et al., Ferraro and Price, Allcott and Kessler, Allcott and Rogers; 2021 Elsevier Inc. is licensed under a All Rights Reserved license
- Figure 25 (Section 4) © P. Wesley Schultz, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein and Vladas Griskevicius; 2007 Association for Psychological Science is licensed under a All Rights Reserved license
- Figure 26 (Section 4) © 2020 Nakayachi, Ozaki, Shibata and Yokoi is licensed under a CC BY (Attribution) license
- Figure 27 (Section 4) © 1969, Springer-Verlag New York Inc. is licensed under a All Rights

Reserved license

- Figure 28 (Section 4) © 2010 Elsevier Ltd. is licensed under a All Rights Reserved license
- Table 3 (Section 4) © 1987, American Psychological Association is licensed under a All Rights Reserved license
- Figure 29 (Section 4) © Roberto A. Weber and Colin F. Camerer; 2003 INFORMS is licensed under a All Rights Reserved license
- Figure 32 (Section 4) © Murakami et al. 2021 is licensed under a CC BY-NC (Attribution NonCommercial) license
- Figure 33 (Section 4) © 1997 Elsevier Science Inc. is licensed under a All Rights Reserved license
- Table 4 (Section 4) © 1997 Elsevier Science Inc. is licensed under a All Rights Reserved license
- Figure 34 (Section 4) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 5 (Section 4) © Michael Conlin, Ted O'Donoghue and Timothy J. Vogelsang; 2007 American Economic Association is licensed under a All Rights Reserved license
- Figure 35 (Section 4) © Minjung Koo and Ayelet Fishbach; 2012 by JOURNAL OF CONSUMER RESEARCH, Inc. is licensed under a All Rights Reserved license
- Table 6 (Section 4) © 2001 by David Genesove and Christopher Mayer is licensed under a All Rights Reserved license
- Table 7 (Section 4) © Thierry Post, Martijn J. van den Assem, Guido Baltussen and Richard H. Thaler; 2008 American Economic Association is licensed under a All Rights Reserved license
- Table 8 (Section 4) © Brendan Nyhan and Jason Reifler; 2010 Springer is licensed under a All Rights Reserved license
- Table 9 (Section 4) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Figure 36 (Section 4) © Nina Mazar, On Amir and Dan Ariely; Journal of Marketing Research adapted by Arthur Caplan is licensed under a All Rights Reserved license

REFERENCES

Abbink, K., Irlenbusch, B., and Renner, E. (2006) Group size and social ties in microfinance institutions. *Economic Inquiry* 44(4), 614-628.

Ainslie, G.W. (1992) *Picoeconomics*. Cambridge, UK: Cambridge University Press.

Akbulut-Yuksel, M. and Boulatoff, C. (2021) The effects of a green nudge on municipal solid waste: evidence from a clear bag policy. *Journal of Environmental Management and Economics* 106, 102404.

Allcott, H. (2011) Social norms and energy conservation. *Journal of Public Economics* 95(9), 1082-1095.

Alm, J., Cherry, T., Jones, M., and McKee, M. (2010) Taxpayer information assistance services and tax compliance behavior. *Journal of Economic Psychology* 31(4), 577-586.

Alper, M., Durose, M.R., and Markman, J. (2018) 2018 update on prisoner recidivism: a 9-Year follow-up period (2005-2014). Special Report. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. Retrieved from the internet on October 5, 2020 at https://www.bjs.gov/content/pub/pdf/18upr9yfup0514.pdf.

American Gaming Association (2019) *American Attitudes on Casino Gaming*. Retrieved from the internet on May 6, 2020 at https://www.americangaming.org/resources/american-attitudes-on-casino-gaming-2019/.

American Lung Association (2020) *Tobacco Use Among Children and Teens*. Retrieved from the internet on November 16, 2020 at https://www.lung.org/quit-smoking/smoking-facts/tobacco-use-among-children.

Andreoni, J. and Miller, J.H. (1993) Rational cooperation in the finitely repeated Prisoner's Dilemma: experimental evidence. *The Economic Journal* 103, 570-585.

Angner, E. and Loewenstein, G. (2012) Behavioral economics. *The Handbook of the Philosophy of Science: Philosophy of Economics*, pp. 641-690. Elsevier.

Ariely, D. (2008) *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York, NY: HarperCollins Publishing.

Ariely, D. and Wertenbroch, K. (2002) Procrastination, deadlines, and performance: self-control by precommitment. *Psychological Science* 13(3), 219-224.

Ariely, D., Gneezy, U., and Haruvy, E. (2018) Social norms and the price of zero. *Journal of Consumer Psychology, Special Issue: Marketplace Morality* 28(2), 180-191.

Ariely, D., Loewenstein, G., and Prelec, D. (2003) Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics* 118, 73-105.

Arkes, H.R., Joyner, C.A., Pezzo, M.V., Nash, J.G., Siegel-Jacobs, K., and Stone, E. (1994) The psychology of windfall gains. *Organizational Behavior and Human Decision Processes* 59, 331-347.

Asch, S.E. (1951) Effects of group pressure on the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, Leadership and Men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.

Ashraf, M., Karlan, D., and Yin, W. (2006) Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics* 121(2), 635-672.

Associated Press (2019) Re-purposed parking meters in Steamboat Springs let hikers donate to trail maintenance on the spot. Retrieved from the internet on May 3, 2020 at https://theknow.denverpost.com/2019/07/01/parking-meters-donate-trail-maintenance-steamboat-springs/218245/.

Bahns, A.J., Pickett, K.M., and Crandall, C.S. (2011) Social ecology of similarity: big schools, small schools and social relationships. *Group Processes and Intergroup Relations* 15(1), 119-131.

Bañas-Garza, P., Capraro, V., and Rascón-Ramírez, E. (2018) Gender difference in altruism on Mechanical Turk: expectations and actual behavior. *Economics Letters* 170, 19-23.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C. (2015) The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics* 7(1), 22-53.

Banerjee, A.V. (2013) Microcredit under the microscope: what have we learned in the past two decades, and what do we need to know? *Annual Review of Economics* 5, 487-519.

Barrera-Osorio, F., Bertrand, M., Linden, L.L., and Perez-Calle, F. (2008) Conditional cash transfers in education: design features, peer and sibling effects evidence from a randomized experiment in Colombia. Policy Research Working Paper 4580. Washington, DC: The World Bank, Human Development Network, Education Team.

Bateson, M., Nettle, D., and Roberts, G. (2006) Cues of being watched enhance cooperation in a real-world setting. *Biology Letters* 2(3), 412-414.

Bauer, M.A., Wilkie, J.E.B., Kim, J.K., and Bodenhausen, G.V. (2012) Cuing consumerism: situational materialism undermines personal and social well-being. *Psychological Science* 23, 517-523.

Beard, T.R. and Beil, R. (1994) Do people rely on the self-interested maximization of others? An experimental test. *Management Science* 40, 252-262.

Becker, G.S. and Murphy, K.M. (1988) A theory of rational addiction. *Journal of Political Economy* 96, 675-700.

Bellemare, C., Gousse, M., Lacroix, G., and Marchand, S. (2023) Physical disability and labor market discrimination: evidence from a video resume field experiment. *American Economic Journal: Applied Economics* 15(4), 452-476.

Belloc, M., Bilancini, E., Boncinelli, L., and D'Alessandro, S. (2019) Intuition and deliberation in the stag hunt game. *Scientific Reports 9*, Article number: 14833. Retrieved from the internet on May 5, 2020 at https://www.nature.com/articles/s41598-019-50556-8.

Benartzi, S. and Thaler, R.H. (1995) Myopic loss aversion and the equity premium puzzle. The Quarterly *Journal of Economics* 110(1), 73-92.

Benzion, U., Rapoport, A., and Yagil, J. (1989) Discount rates inferred from decisions: an experimental study. *Management Science* 35(3), 270-284.

Berg, J.E., Dickhaut, J., and McCabe, K. (1995) Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122-142.

Bergen, B.K. (2016) *What the F: What Swearing Reveals About Our Language, Our Brains, and Ourselves*. New York, NY: Basic Books.

Berger, J. and Pope, D. (2011) Can losing lead to winning? *Management Science* 57(5), 817-827.

Berger, J., Meredith, M., and Wheeler, S.C. (2008) Contextual priming: where people vote affects how people vote. *Proceedings of the National Academy of Sciences (PNAS)* 105(26), 8846-8849.

Berninghaus, S.K., Erhart, K-M and Keser, C. (2002) Conventions and local interaction structures: Experimental evidence. *Games and Economic Behavior* 39, 177-205.

Bertrand, M. and Mullainathan, S. (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94(4), 991-1013.

Bevan, G. and Hamblin, R. (2009) Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK. *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 172(1), 161-190.

Bevc, C.A., Retrum, J.H., and Varda, D.M. (2015) New perspectives on the "silo effect": initial comparisons of network structures across public health collaboratives. *American Journal of Public Heath* 105(2), S230-S235.

Bickel, W.K., Odum, A.L., and Madden, G.J. (1999) Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology* 146, 447-454.

Biederman, I. (1972) Perceiving real-world scenes. *Science* 177, 77-80.

Binmore, K., Proulx, C., Samuelson, L., and Swierzbinski, J. (1998) Hard bargains and lost opportunities. *Economic Journal* 108,1279-1298.

Blau, F.D. and Kahn, L.M. (2017) The gender wage gap: extent, trends, and explanations. *Journal of Economic Literature* 55(3), 789-865.

Bolton, G.E. and Zwick,R. (1995) Anonymity versus punishment in ultimatum bargaining. *Games and Economic Behavior* 10, 95-121.

Brooks, D. (2021) Why is it ok to be mean to the ugly? *New York Times*, June 24, 2001. Retrieved from the internet on May 25, 2021 at https://www.nytimes.com/2021/06/24/opinion/why-is-it-ok-to-be-mean-to-the-ugly.html.

Buchan, N.R., Croson, R.T.A., and Dawes, R.M. (2000) Who's with me? Direct and indirect trust and reciprocity in China, Japan, Korea, and the United States. University of Wisconsin working paper.

Buchan, N.R., Croson, R.T.A., and Johnson, E.J. (2004) When do fair beliefs influence bargaining behavior? Experimental bargaining in Japan and the United States. *Journal of Consumer Research* 31(1), 181-190.

Busenitz, L.W. and Barney, J.B. (1997) Differences between entrepreneurs and managers in large organizations: biases and heuristics in strategic decision-making. *Journal Business Venturing* 12, 9-30.

Camerer, C.F. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*. New York, NY: Russell Sage Foundation and Princeton, N.J.: Princeton University Press.

Camerer, C.F. and Hogarth, R.M. (1999) The effects of financial incentives in economics experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* 18, 7-42.

Camerer, C.F. and Knez, M. (1994) Creating "expectational assets" in the laboratory: "Weakest-link" coordination games. *Strategic Management Journal* 15, 101-119.

Camerer, C.F. and Loewenstein, G. (1993) Information, fairness, and efficiency in bargaining. In B. Mellers and J. Baron (Eds.), *Psychological Perspectives on Justice: Theory and Applications*. Cambridge, U.K.: Cambridge University Press, 155-179.

Camerer, C. and Loewenstein, G. (2004) Behavioral economics: past, present, future. In C. Camerer, G. Loewenstein, & M. Rabin (Eds.), *Advances in Behavioral Economics* (pp. 3-51). Princeton, NJ: Princeton University Press.

Camerer, C. and Lovallo, D. (1999) Overconfidence and excess entry: an experimental approach. *American Economic Review* 89(1), 306-318.

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997) Labor supply of New York City cab drivers: one day at a time. *Quarterly Journal of Economics* 112(2), 407-441.

Cameron, L.A. (1999) Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry* 27, 47-59.

Cannon, W. B. (1942) "Voodoo" death. *American Anthropologist* 44, 182–190.

Caplan, A.J. and Gilbert, J. (2008) 'D' is for dillydally? *Applied Economics Letters* 15(14), 1085-1088.

Capra, M.C., Goeree, J.K., Gomez, R., and Holt C. (1999) Anomalous behavior in a travelers' dilemma? *American Economic Review* 89, 678-690.

Carter, M.R. and Castillo, M. (2004) Morals, markets and mutual insurance: using economic experiments to study recovery from Hurricane Mitch. In C.B. Barrett (ed.). *Exploring the Moral Dimensions of Economic Behavior*. London: Routledge, 268–287.

Carter, M.R., and Castillo, M. (2011) Trustworthiness and social capital in South Africa: analysis of actual living standards data and artifactual field experiments. *Economic Development and Cultural Change* 59(4), 695–722.

Cartwright, E. (2018) *Behavioral Economics*, Third Edition. New York, NY: Routledge.

Chan, K.S., Godby, R., Mestelman, S., and Muller, R.A. (2002) Crowding out voluntary contributions to public goods. *Journal of Economic Behavior and Organization* 48, 305-317.

Cheek, N.N., Blackman, S.F., and Pronin, E. (2020) Seeing the subjective as objective: people perceive the taste of those they disagree with as biased and wrong. *Journal of Behavioral Decision Making*. DOI: 10.1002/bdm.2201.

Chiang, A.C. and Wainwright, K. (2005) *Fundamental Methods of Mathematical Economics*, Fourth Edition. New York, NY: McGraw-Hill Higher Education.

Cialdini, R.B., Demaine, L.J., Sagarin, B.J., Barrett, D.W., Kelton, R., and Winter, P.L. (2006) Managing social norms for persuasive impact. *Social Influence* 1(1), 3-15.

Cialdini, R.B., Reno, R.R., and Kallgren, C.A. (1990) A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places." *Journal of Personality and Social Psychology* 58(6), 1015-1026.

City of Denver, CO (2015) Denver's ten-year plan to end homelessness 2005-2015. A Report to Mayor Michael B. Hancock by Denver's Road Home. Retrieved from the internet on May 12, 2020 at https://www.denvergov.org/content/dam/denvergov/Portals/646/documents/Zoning/text_amendments/Group_Living/Denvers_Road_Home_Ten_Year_Plan_to_End_Homelessness.pdf.

Clotfelter, C. and Cook, P. (1991) Lotteries in the real world. *Journal of Risk and Uncertainty* 4, 227-232.

Cohen, A.R. (1957) Need for cognition and order of communication as determinants of opinion change. In C. I. Hovland et al. (Eds.), *The Order of Presentation in Persuasion*. New Haven: Yale University. Press, pp. 79-97.

Cohen-Zada, D., Krumer, A., and Shtudiner, Z. (2017) Psychological momentum and gender. *Journal of Economic Behavior and Organization* 135, 66-81.

Coleman, S. (1996) The Minnesota income tax compliance experiment: state tax results. Minnesota Department of Revenue.

Collins, R. (1989) *The Arab Conquest of Spain 710–797*. Oxford, UK / Cambridge, USA: Blackwell Publishers.

Collins, R. and Sherstyuk, K. (2000) Spatial competition with three firms: An experimental study. *Economic Inquiry* 38, 73-94.

Conlin, M., O'Donoghue, T., and Vogelsang, T.J. (2007) Projection bias in catalog orders. *American Economic Review* 97(4), 1217-1249.

Cooper, R., DeJong, D.V., Forsythe, R., and Ross, T.W. (1989) Communication in the battle of the sexes game: some experimental results. *The RAND Journal of Economics* 20(4), 568-587.

Cooper, R., DeJong, D., Forsythe, B., and Ross, T. (1990) Selection criteria in coordination games: Some experimental results. *American Economic Review* 80, 218-233.

Cooper, R., DeJong, D., Forsythe, B., and Ross, T. (1994) Alternative institutions for resolving coordination problems: Experimental evidence on forward induction and preplay communication. In J. Friedman (Ed.), *Problems of Coordination in Economic Activity*. Dordrecht: Kluwer.

Croson, R.T.A. (1996) Information in ultimatum games: An experimental study. *Journal of Economic Behavior and Organization* 30, 197-212.

Croson, R. and Buchan, N. (1999) Gender and culture: international experimental evidence from trust games. *AEA Papers and Proceedings* 89(2), 386-391.

D'Amato, A., Mancinelli, S., Zoli, M., (2016) Complementarity vs substitutability in waste management behaviors. *Ecological Economics* 123, 84-94.

Damon, W. (1980) Patterns of change in children's social reasoning: A two-year longitudinal study. *Child Development* 51, 1010-1017.

Danziger, S., Levav, J., and Avnaim-Pesso, L. (2011) Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences (PNAS)* 108(17), 6889-6892.

Darley, J.M. and Batson, C.D. (1973) "From Jerusalem to Jericho": a study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology 27*(1), 100-108.

Darley, J.M. and Gross, P.H. (1983) A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology 44*(1), 20-33.

Davidai, S., Gilovich, T. and Ross, L. (2012) The meaning of default options for potential organ donors. Proceedings of the National Academy of Sciences, 109(38), 15201-15205.

De Crescenzo, F., et al. (2018) Comparative efficacy and acceptability of psychosocial interventions for individuals with cocaine and amphetamine addiction: a systematic review and network meta-analysis. *PLoS Med* 15(12), e1002715.

DeBruine, L.M. (2002) Facial resemblance enhances trust. *Proceedings of the Royal Society of London* B 269, 1307–1312.

Degenhardt L. and Hall, W. (2012) Extent of illicit drug use and dependence, and their contribution to the global burden of disease. *Lancet* 379(9810), 55-70.

DellaVigna, S. and Malmendier, U. (2006) Paying not to go to the gym. *The American Economic Review* 96(3), 694-719.

Dixit, A.K. and Nalebuff, B.J. (1991) *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*. New York, NY: W.W. Norton & Company.

Dranove, D., Kessler, D., McClellan, M., and Satterthwaite, M. (2003) Is more information better? The effects of "report cards" on health care providers. *Journal of Political Economy* 11(3), 555-588.

Duggan, M. and Levitt, S.D. (2002) Winning isn't everything: corruption in sumo wrestling. *American Economic Review* 92(5), 1594-1605.

DynamiCare Health (2020) Proven to help quit or cut down on substances. Retrieved from the internet on September 15, 2020 at https://dynamicarehealth.com.

Eckel, C. and Grossman, P. (2001) Chivalry and solidarity in ultimatum games. *Economic Inquiry* 30, 171-188.

Ellsberg, D. (1961) Risk, ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75, 643-669.

Employee Benefit Research Institute (EBRI) (2020) *2020 Retirement Confidence Survey Summary*

Report. April 23, 2020. Retrieved from the internet on May 6, 2020 at https://www.ebri.org/docs/default-source/rcs/2020-rcs/2020-rcs-summary-report.pdf?sfvrsn=84bc3d2f_7.

Erat, S. and Gneezy, U. (2012) White lies. *Management Science* 58(4), 723-733.

Euser, P. (2008) *The Laweiplein: Evaluation of the Reconstruction Into a Square With Roundabout*. Noordelijke Hogeschool Leeuwarden, The Netherlands.

Falk, A., Fehr, E., and Fischbacher, U. (2003) On the nature of fair behavior. *Economic Inquiry* 41(1), 20-26.

Fehr, E. and Fischbacher, U. (2004) Third-party punishment and social norms. *Evolution and Human Behavior* 25, 63-87.

Fehr E. and Gächter, S. (2000) Cooperation and punishment in public goods experiments. *American Economic Review* 90, 980-994.

Fehr, E. and Schmidt, K.M. (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114, 817-868.

Fischbacher, U. and Föllmi-Heusi, F. (2013) Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association* 11(3), 525-547.

Fischhoff, B. and Beyth, R. (1975) I knew it would happen: remembered probabilities of once-future things. *Organizational Behavior and Human Performance* 13(1), 1-16.

Forsythe, R., Horowitz, J.L., Savin, N.E., and Sefton, M. (1994) Fairness in simple bargaining experiments. *Games and Economic Behavior* 6, 347-369.

Fryer, Jr, R.G., Levitt, S.D., List, J., Sadoff, S. (2022) Enhancing the efficacy of teacher incentives through framing: a field experiment. *American Economic Journal: Economic Policy* 14(4), 269-299.

Fung, A. and O'Rourke, D. (2000) Reinventing environmental regulation from the grassroots up: explaining and expanding the success of the Toxic Release Inventory. *Environmental Management* 25(2), 115-127.

Gächter, S., Herrmann, B., and Thöni, C. (2010) Culture and cooperation. *Philosophical Transactions of the Royal Society B* – Biological Sciences 365, 2651-2661.

Gaissmaier, W. and Marewski, J.N. (2010) Forecasting elections with mere recognition from small, lousy samples: a comparison of collective recognition, wisdom of crowds, and representative polls. *Judgement and Decision Making* 6(1), 73-88.

Galai, D. and Sade, O. (2006) The "ostrich effect" and the relationship between the liquidity and the yields of financial assets. *Journal of Business*. 79(5), 2741–2759.

Gallup Organization (1993) *The American Public's Attitude References and Notes Toward Organ Donation and Transplantation.* Princeton, NJ: Gallup Organization.

García-Amado, L.R., Ruiz-Pèrez, M., and Barrasa García, S. (2013) Motivation for conservation: assessing integrated conservation and development projects and payments for environmental services in La Sepultura Biosphere Reserve in Chiapas, Mexico. *Ecological Economics* 89, 92-100.

GBD 2016 DALYs and HALE Collaborators (2017) Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Genesove, D. and Mayer, C. (2001) Loss aversion and seller behavior: evidence from the housing market. *The Quarterly Journal of Economics* 116(4), 1233-1260.

Gershenson, S., Hart, C.M.D., Hyman, J., Lindsay, C.A., Papageorge, N.W. (2022) The long-run impacts of same-race teachers. *American Economic Journal: Economic Policy* 14(4), 300-342.

Gigerenzer, G. and Gaissmaier, W. (2011) Heuristic decision making. *Annual Review of Psychology* 62, 451-482.

Gladwell, M. (2002) *The Tipping Point: How Little Things Can Make a Big Difference*. New York, NY: Little, Brown and Company.

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., and Soutter, C.L. (2000) Measuring trust. *Quarterly Journal of Economics* 115, 811–846.

Gneezy, U. and Rustichini, A. (2000) A fine is a price. *The Journal of Legal Studies* 29(1), 1-17.

Goldstein, N.J., Cialdini, R.B., and Griskevicius, V. (2008) A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research* 35, 472-482.

Goodreads.com (2022) *Seneca Quotes*. Retrieved from the internet on May 15, 2021 at https://www.goodreads.com/author/quotes/4918776.Seneca.

Gourville, J.T. and Soman, D. (1998) Payment depreciation: the behavioral effects of temporarily separating payments from consumption. *Journal of Consumer Research* 25(2), 160-174.

Green, L. and Mehr, D.R. (1997) What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice* 45, 219-26.

Greenwald, A.G., Carnot, C.G., Beach R., and Young, B. (1987) Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology* 72(2), 315-318.

Grether, D.M. and Plott, C.R. (1979) Economic theory of choice and the preference reversal phenomenon. *The American Economic Review* 69(4), 623-638.

Hale, G., Regev, T., and Rubinstein, Y. (2021) Do looks matter for an academic career in economics? CEPR Discussion Paper No. DP15893, Available at SSRN: https://ssrn.com/abstract=3805308.

Halevy, Y. (2007) Ellsberg revisited: an experimental study. *Econometrica* 75, 503-536.

Hamermesh, D.S., and Biddle, J.E. (1994) Beauty and the labor market. *American Economic Review* 84, 1174-1194.

Haney, C., Banks, C., and Zimbardo, P. (1973) Interpersonal dynamics in a simulated prison. International *Journal of Criminology and Penology* 1, 69-97.

Hanley, N., Shogren, J., and White, B. (2007) *Environmental Economics: In Theory and Practice*. London, UK: Red Globe Press.

Harari, Y.N. (2015) *Sapiens: A Brief History of Humankind*. New York: Harper Perennial.

Harbaugh, W.T., Krause, K., and Liday, S.J. (2003) Bargaining by children. University of Oregon Economics Working Paper No. 2002-4. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=436504.

Harford, T. (2016) *Messy: The Power of Disorder to Transform Our Lives*. New York, NY: Riverhead Books.

Hart. C. (2013) *High Price: A Neuroscientist's Journey of Self-Discovery That Challenges Everything You Know About Drugs and Society*. New York, NY: HarperCollins Publishers.

Harvey, G.E. (1925) *History of Burma from the Earliest Times to 10 March 1824*: *The Beginning of the English Conquest*. London: Longmans, Green and Co.

Haselhuhn, M.P., Pope, D.G., Schweitzer, M.E., and Fishman, P. (2012) The impact of personal experience on behavior: evidence from video-rental fines. *Management Science* 58(1), 52-61

Hastie, R. and R. Dawes (2001) *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks: Sage Publications.

Hawking, S. (2017) *A Brief History of Time*. New York, NY: Bantam Books.

Heath, C. and Soll, J.B. (1996) Mental budgeting and consumer decisions. *Journal of Consumer Research* 23(1), 40-52.

Heath, C. and Tversky, A. (1991) Preference and belief: ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4, 5-28.

Heffetz, O., O'Donoghue, T., Schneider, H.S. (2022) Reminders work, but for whom? Evidence from New York City parking ticket recipients. *American Economic Journal: Economic Policy* 14(4), 343-370.

Henrich, J. (2000) Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *American Economic Review* 90, 973-979.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N. S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F. W., Patton, J. Q., & Tracer, D. (2005) Economic man in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, *28*(6), 795-815.

Henrich, J., Boyd, R., Bowles S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001) In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review* 91(2), 73-78.

Heukelom, F. (2006) Kahneman and Tversky and the origins of behavioral economics. Tinbergen Institute Discussion Paper TI 2007-003/1.

Heurer, L. and Orland, A. (2019) Cooperation in the Prisoner's Dilemma: an experimental comparison between pure and mixed strategies. Royal Society Open Science 6(7). Retrieved from the internet on May 4, 2019 at https://doi.org/10.1098/rsos.182142.

Heyman, J. and Ariely, D. (2004) Effort for payment: a tale of two markets. *Psychological Science* 15(11), 787-794.

Hindu.com (2018) What is sure thing principle in decision theory? Retrieved from the Internet on September 10, 2020 at https://www.thehindu.com/opinion/op-ed/what-is-sure-thing-principle-in-decision-theory/article25664838.ece.

History.com (2020) *First Battle of The Sexes*. Retrieved from the internet on May 7, 2020 at https://www.history.com/this-day-in-history/first-battle-of-the-sexes.

Hoffman, E., McCabe, K., and Smith, V.L. (1996) On expectations and monetary stakes in ultimatum games. *International Journal of Game Theory* 25, 289-301.

Hoffman, E., McCabe, K., Shachat, K., and Smith, V.L. (1994) Preferences, property rights and anonymity in bargaining games. *Games and Economic Behavior* 7, 346-380.

Hogarth, R.M. and Einhorn, H.J. (1992) Order effects in belief updating: the belief-adjustment model. *Cognitive Psychology* 24, 1-55.

Homonoff, T. (2018) Can small incentives have large effects? The impact of taxes vs. bonuses on disposable bag use. *American Economic Journal: Economic Policy* 10(4), 177-210.

Huck, S. (1999) Responder behavior in ultimatum offer games with incomplete information. *Journal of Economic Psychology* 20, 183-206.

Huddleston, C. (2019) 58% of Americans have less than $1,000 in savings, survey finds. *GOBankingRates*. Retrieved from the internet on May 6, 2020 at https://finance.yahoo.com/news/58-americans-less-1-000-090000503.html.

Jepsen, R. (2019) Panhandling: see how some cities are taking an innovative approach to deter panhandling while promoting donations to bolster services for homeless populations. Alliance for Innovation. Retrieved from the internet on May 7, 2020 at https://www.transformgov.org/articles/%E2%80%9Cdonation-meters%E2%80%9D-address-homelessness-and-panhandling-0.

Jessoe, K., Lade, G.E., Loge, F., and Spang, E. (2020) Spillovers from behavioral interventions: experimental evidence from water and energy use. *Journal of the Association of Environmental and Resource Economists*. https://doi.org/10.1086/711025.

Johansson-Stenman, O., Mahmud, M., and Martinsson, P. (2009) Trust and religion: experimental evidence from rural Bangladesh. *Economica* 76, 462–485.

Johnson, E.J. and Goldstein, D. (2003) Do defaults save lives? *Science* 302, 1338-1339.

Just, D.R. (2013) *Introduction to Behavioral Economics*. Oak Brook, Il: Wiley Global Education.

Just, D.R. and Wansink, B. (2011) The flat-rate pricing paradox: conflicting effects of all-you-can-eat buffet pricing. *The Review of Economics and Statistics* 93(1), 193-200.

Kahneman, D. (2011) *Thinking, Fast and Slow*. New York, NY: Macmillan Publishers.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2), 263-292.

Kahneman, D. and Tversky, A. (1984) Choices, values, and frames. *American Psychologist* 39(4), 341-350.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1986a) Fairness and the assumptions of economics. *Journal of Business* 59(4), Part 2: The Behavioral Foundations of Economic Theory, S285-S300.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1986b) Fairness as a constraint on profit seeking: entitlements in the market. *American Economic Review* 76(4), 728-741.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1990) Experimental tests of the endowment effect and the Coase Theorem. *Journal of Political Economy*, 98, 1325-1348.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1991) The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives* 5(1), 193-206.

Kangas, O., Jauhiainen, S., Simanainen, M., and Ylikännö, M. (2019) *The Basic Income Experiment 2017–2018 in Finland. Preliminary results*. Ministry of Social Affairs and Health, Helsinki. Retrieved from the internet on September 23, 2020 at https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161361/Report_The%20Basic%20Income%20Experiment%2020172018%20in%20Finland.pdf.

Karabenick, S.A. and Addy, M.M. (1979) Locus of control and sex differences in skill and chance risk-taking conditions. *Journal of General Psychology*, 100(2), 215–228.

Karlsson, N., Loewenstein, G., and Seppi, D. (2009) The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty* 38(2), 95–115.

Kaza, S., Yao, L. C., Bhada-Tata, P., Van Woerden, F. (2018) *What a Waste 2.0 : A Global Snapshot of Solid Waste Management to 2050. Urban Development*. Washington, DC: World Bank.

Kela (2020) Results of Finland's basic income experiment: small employment effects, better perceived economic security and mental well-being. Retrieved from the internet on September 23, 2020 at https://www.kela.fi/web/en/news-archive/-/asset_publisher/lN08GY2nIrZo/content/results-of-the-basic-income-experiment-small-employment-effects-better-perceived-economic-security-and-mental-wellbeing#:~:text=In%20the%20basic%20income%20experiment,were%20actively%20looking%20for%20work.&text=The%20experiment%20was%20implemented%20by,Social%20Insurance%20Institution%20of%20Finland).

Kelley, E.J. (1958) The importance of convenience in consumer purchasing. *Journal of Marketing* 23(1), 32-38.

Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*. UK: Palgrave Macmillan.

Kim, J. (2023) The effects of time preferences on cooperation: experimental evidence from infinitely repeated games. *American Economic Journal: Microeconomics* 15(1), 618-37.

Kirschenbaum, D.S., Humphrey, L.L., and Malett, S.D. (1981) Specificity of planning in adult self-control: an applied investigation. *Journal of Personality and Social Psychology* 40(5), 941-950.

Knack, S. and Keefer, P. (1997) Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics* 112(4), 1251-1288.

Knez, M.J., and Camerer, C.F. (1995) Social comparison and outside options in 3-person ultimatum games. *Games and Economic Behavior* 10, 165-194.

Knight, C. and Haslem, S.A. (2010) The relative merits of lean, enriched, and empowered offices: an experimental examination of the impact of workspace management strategies on well-being and productivity. *Journal of Experimental Psychology: Applied* 16(2), 158-172.

Koo, M. and Fishbach, A. (2012) The small-area hypothesis: effects of progress monitoring on goal adherence. *Journal of Consumer Research* 39, 493-509.

Kreps, D. M., Milgrom, P., Roberts, J. and Wilson, R. (I982) Rational cooperation in the finitely repeated Prisoners' Dilemma. *Journal of Economic Theory* 27, 245-252.

Laibson, D. (1979) Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 112(2), 443-477.

Landy, D. and Sigall, H. (1974) Beauty is talent: task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology* 29, 299-304.

Larrick, R.P. and Blount, S. (1997) The claiming effect: Why players are more generous in social dilemmas than in ultimatum games. *Journal of Personality and Social Psychology* 72, 810-825.

Lee, L., Frederick, S., and Ariely D. (2006) Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science* 17(12), 1054-1058.

Leventhal, H., Singer, R., and Jones, S. (1965) Effects of fear and specificity of recommendation upon attitudes and behavior. *Journal of Personality and Social Psychology* 2(1), 20-29.

Levitt, S.D. (1998) Why do increased arrest rates appear to reduce crime: deterrence, incapacitation, or measurement error?" *Economic Inquiry* 36, 353-372.

Levitt, S.D. (2004) Testing theories of discrimination: evidence from Weakest Link". *Journal of Law and Economics* 48, 431-452.

Levitt, S.D. and Dubner, S.J. (2005) *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York, NY: HarperCollins.

Levitt, S.D., List, J.A., Neckermann, S., and Sadoff, S. (2016) The behavioralist goes to school: leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy* 8(4), 183-219.

Lichtenstein, S. and Slovic, P. (1971) Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*(1), 46-55.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., and Combs, B. (1978) Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory* 4, 551-578.

Linos, E., Prohofsky, A., Ramesh, A., Rothstein, J., and Unrath, M. (2022) Can nudges increase take-up of the EITC?: evidence from multiple field experiments. *American Economic Journal: Economic Policy* 14(4), 432-52.

List, J.A. (2004) Neoclassical theory versus prospect theory: evidence from the marketplace. *Econometrica* 72(2), 615-625.

List, J.A. (2006) The behavioralist meets the market: measuring social preferences and reputation effects in actual transactions. *Journal of Political Economy* 114(1), 1-37.

List, J.A. and Cherry, T.L. (2000) Learning to accept in ultimatum games: Evidence from an experimental design that generates low offers. *Experimental Economics* 3, 11-31.

List, J.A. and Shogren, J.F. (1998) The deadweight loss of Christmas: comment. *American Economic Review* 88(5), 1350-1355.

Littlewood, J. E. (1953) *A Mathematician's Miscellany*. London: Meuthen & Co. Ltd.

Loewenstein, G. (1987) Anticipation and the valuation of delayed consumption. *The Economic Journal* 97(387), 666-684.

Loewenstein, G. (1999) Experimental economics from the vantage-point of behavioural economics. *The Economic Journal* 109, F25-F34.

Loewenstein, G. (2005) Hot-cold empathy gaps and medical decision making. *Health Psychology* 24(4), S49-S56.

Loewenstein, G. and Prelec, D. (1992) Anomalies in intertemporal choice: evidence and an interpretation. Anomalies in intertemporal choice: evidence and an interpretation. *The Quarterly Journal of Economics* 107(2), 573-597.

Loewenstien, G., O'Donoghue, T. and Rabin, M. (2003) Projection bias in predicting future utility. *Quarterly Journal of Economics* 118(4), 1209-1248.

Loomes, G. and Sugden, R. (1982) Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal* 92(368), 805-824.

Loomes, G. and Sugden, R. (1987) Some implications of a more general form of regret theory. *Journal of Economic Theory* 41, 270-287.

Loomes, G., Starmer, C., and Sugden, R. (1991) Observing violations of transitivity by experimental methods. *Econometrica* 59(2), 425-439.

Madden G.J., Petry, N.M., Badger, G.J., and Bickel, W.K. (1997) Impulsive and self-control choices in opioid-dependent patients and nondrug-using control participants: drug and monetary rewards. *Experimental Clinical Psychopharmacology* 5, 256-262.

Marchese, L. (2016) How the 'silo effect' is hurting cross team collaboration. Retrieved from the internet on December 2, 2020 at https://blog.trello.com/tips-to-improve-cross-team-collaboration.

Martin, S.J., Goldstein, N., and Cialdini, R. (2014) *The B!G Small: Small Changes that Spark Big Influence*. New York, NY: Grand Central Publishing.

Mas-colell, A., Whinston, M.D., and Green, J.R. (1995) *Microeconomic Theory*. Oxford, UK: Oxford University Press.

Mazar, N., On, A., and Ariely, D. (2008) The dishonesty of honest people: a theory of self-concept maintenance. *Journal of Marketing Research* 45(6), 633-644.

McClure, S.M., Li, J., Tomlin, D., Cypert, K.S., Montague, L.M., and Montague, P.R. (2004) Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379-387.

McGrath, R. G., MacMillan, I.,C., and Scheinberg, S. (1992) Elitists, risk-takers, and rugged individualists? An exploratory analysis of cultural differences between entrepreneurs and non-entrepreneurs. *Journal of Business Venturing* 7, 115-135.

McKelvey, R.D. and Palfrey, T.R. (1992) An experimental study of the centipede game. *Econometrica* 60(4), 803-836.

McLeod, S. (2018) Solomon Asch – conformity experiment. *Simply Psychology*. Retrieved from the internet on May 7, 2019 at https://www.simplypsychology.org/asch-conformity.html.

Mehra, R. and Prescott, E.C. (1985) The equity premium puzzle. *Journal of Monetary Economics* 15, 145-461.

Mehta, J., Starmer C., and Sugden. R. (1992) An experimental investigation of focal points in coordination and bargaining: Some preliminary results. In J. Geweke (Ed.), *Decision Making under Risk and Uncertainty: New Models and Findings*. Norwell, Mass.: Kluwer, 211-220.

Mia, M.A., Lee, H-A, Chandran, V.G.R., Rasiah, R., and Rahman, M. (2017) History of microfinance in Bangladesh: a life-cyle theory approach. *Business History* 61(4), 703-733.

Milgram, S. (1963) Behavioral study of obedience. *Journal of Abnormal and Social Psychology* 67, 371-378.

Miravete, E. (2003) Choosing the wrong calling plan? Ignorance and learning. *American Economic Review* 93, 297-310.

Mitzkewitz, M. and Nagel, R. (1993) Experimental results on ultimatum games with incomplete information. *International Journal of Came Theory* 22, 171-198.

Montoy, J.C.C., Dow, W.H., and Kaplan, B.C. (2016) Patient choice in opt-in, active choice, and opt-out HIV screening: randomized clinical trial. *The British Medical Journal* (*BMJ*) 352:h6895. Retrieved from the internet on April 28, 2021 at http://dx.doi.org/10.1136/bmj.h6895.

Morton, H. (2015) Prize-linked savings accounts: offering incentives to save. National Conference of State Legislatures. Retrieved from the internet on June 6, 2020 at https://www.ncsl.org/research/financial-services-and-commerce/prize-linked-savings-accounts-offering-incentives-to-save.aspx.

Murakami, H., Feliciani, C., Nishiyama, Y., and Nishinari, K. (2021) Mutual anticipation can contribute to self-organization in human crowds. *Science Advances* 7:eabe7758.

Murnighan, J.K. and Saxon, M.S. (1998) Ultimatum bargaining by children and adults. *Journal of Economic Psychology*, 19, 415-445.

Myers, D.G. (2010) *Social Psychology*, *Tenth Edition*. New York, NY: McGraw-Hill.

Nakayachi, K., Ozaki, T., Shibakta Y., and Yokoi, R. (2020) Why do Japanese people use masks against COVID-19, even though masks are unlikely to offer protection from infection? *Frontiers in Psychology* 11:1918.

Nash, J. (1950a) Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* 36(1), 48-49.

Nash, J. (1950b) The bargaining problem. *Econometrica* 18(2), 155-162.

Nash, J. (1951) Non-cooperative games. *Annals of Mathematics* 54(2), 286-295.

Niederle, M. and Vesterlund, L. (2007) Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122, 1067-1101.

Nisbett, R.E. and Wilson, T.D. (1977) The halo effect: evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology* 35(4), 250-256.

Northcraft, G.B. and Neale, M.A. (1987) Experts, amateurs, and real estate: an anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes* 39(1), 84-97.

NRF (2020) Retail shrink totaled $61.7 billion in 2019 amid rising employee theft and shoplifting/ORC. Retrieved from the internet on January 31, 2021 at https://nrf.com/media-center/press-releases/retail-shrink-totaled-617-billion-2019-amid-rising-employee-theft-and.

Nyhan, B. and Reifler, J. (2010) When corrections fail: the persistence of political misperceptions. *Political Behavior* 32(2), 303-330.

Odean, T. (1998) Are investors reluctant to realize losses? *Journal of Finance* 52(5), 1775-1789.

OECD (2019) Public social spending is high in many OECD countries. Social Expenditure Update 2019. Retrieved from the internet on September 18, 2020 at https://www.oecd.org/social/soc/OECD2019-Social-Expenditure-Update.pdf.

Ortmann, A., Gigerenzer, G., Borges, B., and Goldstein, D.G. (2008) The recognition heuristic: a fast and frugal way to investment choice? In *Handbook of Experimental Economics Results: Vol. 1* (Handbooks in Economics No. 28), ed. C.R. Plott and V.L. Smith, pp. 993–1003. Amsterdam: North Holland.

Oskamp, S. (1965) Overconfidence in case-study judgements. *Journal of Consulting Psychology* 29(3), 261-265.

Palmer, S.E. (1975) The effects of contextual scenes on the identification of objects. *Memory and Cognition* 3, 519-526.

Pennisi, E. (2018) 2018 Breakthrough of the year: development cell by cell. Science. Retrieved from the internet on July 19, 2019 at https://vis.sciencemag.org/breakthrough2018/.

Perkins, H.W., Linkenbach, J.W., Lewis, M.J., and Neighbors, C. (2010) Effectiveness of social norms media marketing in reducing drinking and driving: a statewide campaign. *Addictive Behaviors* 35, 866-874.

Persky, J. (1995) The ethology of Homo economicus. *Journal of Economic Perspectives* 9(2), 221-231.

Pew Research Center (2009) *Top Stories of 2009: Economy, Obama, Health Care, Summary Findings*. Retrieved from the internet on January 4, 2021 at https://www.pewresearch.org/politics/2009/12/29/top-stories-of-2009-economy-obama-and-health-care/.

Poe, G.L., Clark, J.E., Rondeau, D., and Schulze, W.D. (2002) Provision point mechanisms and field validity tests of contingent valuation. *Environmental and Resource Economics* 23, 105-131.

Pop-Eleches, C., Thirumurthy, H., Habyarimana, J.P., Zivin, J.G., Goldstein, M.P., de Walque, D., MacKeen, L., Haberer, J., Kimaiyo, S., Sidle, J., Ngare, D., and Bangsberg, D.R. (2001) Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS* 25(6), 825-834.

Pope, D.G. and Schweitzer, M.E. (2011) Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *American Economic Review* 101(1), 129-157.

Pope, D.G. and Sydnor, J.R. (2011) What's in a picture? Evidence of discrimination from Prosper.Com. *Journal of Human Resources* 46(1), 53-92.

Pope, D.G., Price, J., and Wolfers, J. (2018) Awareness reduces racial bias. *Management Science* 64(11), 4967-5460.

Post, T., van den Assem, M.J., Baltussen, G., and Thaler, R.H. (2008) 'Deal or No Deal'? Decision making under risk in a large-payoff game show. *American Economic Review* 98, 38-71.

Poundstone, W. (1992) *Prisoner's Dilemma*. New York, NY: Anchor Books.

Prelec, D. and Loewenstein, G. (1998) The red and the black: mental accounting of savings and debt. *Management Science* 17(1), 4-28.

Pronin, E., Lin, D.Y., and Ross, L. (2002) The bias blind spot: perceptions of bias in self versus others. *Personality and Social Psychology Bulletin* 28(3), 369-381.

Pronin, E., Wegner, D.M., McCarthy, K., and Rodriguez, S. (2006) Everyday magical powers: the role of apparent mental causation in the overestimation of personal influence. *Journal of Personality and Social Psychology* 91(2), 218-231.

Rapoport, A., Sundali, J.A., and Potter, R.E. (1996) Ultimatums in two-person bargaining with one-sided uncertainty: Offer games. *International Journal of Game Theory* 25, 475-494.

Raworth, K. (2017) *Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist*. White River Junction, VT: Chelsea Green Publishing, Inc.

Read, D. and Loewenstein, G. (1995) Diversification bias: explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied 1(1), 34-49.*

Read, D. and Loewenstein, G. (1999) Enduring pain for money: decisions based on the perception and memory of pain. *Journal of Behavioral Decision Making* 12, 1-17.

Read, D., Loewenstein, G., and Kalyanaraman, S. (1999a) Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic. *Journal of Behavioral Decision Making, 12,* 257-273.

Read, D., Loewenstein, G., and Rabin, M. (1999b) Choice bracketing. *Journal of Risk and Uncertainty* 19(1), 171-197.

Redelmeier, D.A., Katz, J., and Kahneman, D. (2003) Memories of colonoscopy: a randomized trial. *Pain* 104(1-2), 187-194.

Rondeau, D., Schulze, W.D., and Poe, G.L. (1999) Voluntary revelation of the demand for public goods using a provision point mechanism. *Journal of Public Economics* 72, 455-470.

Rose, S.K., Clark, J., Poe, G.L., Rondeau, D., and Schulze, W.D. (2002) The private provision of public goods: tests of a provision point mechanism for funding green power programs. *Resource and Energy Economics* 24, 131-155.

Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., and Zamir,S. (1991) Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An experimental study. *American Economic Review* 81, 1068-1095.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., and Birrell, P. (1978) From individual to group impressions: availability heuristics in stereotype formation. *Journal of Experimental Social Psychology* 14, 237-255.

Rudder, C. (2014) We experiment on human beings! Oktrends: Dating Research From OkCupid. Retrieved from the internet on December 8, 2020 at https://www.gwern.net/docs/psychology/okcupid/weexperimentonhumanbeings.html.

Samson, A. (2012) Do you diversify? Research shows that variety seeking isn't always the best strategy. *Psychology Today*. Retrieved from the internet on October 15, 2020 at https://www.psychologytoday.com/us/blog/consumed/201203/do-you-diversify.

Samson, A. (2019) An Introduction to Behavioral Economics. Retrieved from the internet on January 17, 2019 at https://www.behavioraleconomics.com/resources/introduction-behavioral-economics/.

Scharlemann, J.P.W., Eckel, C.C., Kalcelnik, A., and Wilson, R.K. (2001) The value of a smile: game theory with a human face. *Journal of Economic Psychology* 22, 617–640.

Schelling, T.C. (1989) Strategy and self-command. *Negotiation Journal* 5(4), 343-347.

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., and Griskevicius, V. (2007) The constructive, destructive, and reconstructive power of social norms. *Psychological Science* 18(5), 429-434.

Shampanier, K., Mazar, N., and Ariely, D. (2007) Zero as a special price: the true value of free products. *Marketing Science* 26(6), 742-757.

Shapero, A. (1975) The displaced, uncomfortable entrepreneur. *Psychology Today* 9(6), 83-88.

Shefrin, H.M. and Thaler, R.H. (1988) The behavioral life cycle hypothesis. *Economic Inquiry* 26(4), 609–643.

Shin, J. and Ariely, D. (2004) Keeping doors open: the effect of unavailability on incentives to keep options viable. *Management Science* 50(5), 575-586.

Simonsohn, U. and Loewenstein, G. (2006) Mistake #37: the effect of previously encountered prices on current housing demand. *Economic Journal* 116, 175-199.

Simonson, I. (1990) The effect of purchase quantity and timing on variety seeking behavior. *Journal of Marketing Research* 32, 150-162.

Simonson, L. and Winer, R. S. (1992) The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research* 19, 133-138.

Skyrms, B. (2004) *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.

Slonim, R., and Roth, A.E. (1998) Learning in high stakes ultimatum games: an experiment in the Slovak Republic. *Econometrica* 63(3), 569–596.

Snook B., Zito, M., Bennell, C., and Taylor, P.J. (2005) On the complexity and accuracy of geographic profiling strategies. *Journal of Quantitative Criminology* 21, 1-26.

Solnick, S. (2001) Gender differences in the ultimatum game. *Economic Inquiry* 39, 189-200.

Solnick, S. and Hemenway, D. The deadweight loss of Christmas: comment. *American Economic Review* 86(5), 1299-1305.

Spaniel, W. (2011) *Game Theory 101: The Complete Textbook*. CreateSpace Independent Publishing Platform.

Stanley, D.A., Sokol-Hessner, P., Banaji, M.R., and Phelps, E.A. (2011) Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academies of Science (PNAS)* 108, 7710–7715.

Stephens, R. and Robertson, O. (2020) Swearing as a response to pain: assessing hypoalgesic effects of novel "swear" words. *Frontiers in Psychology* 11, Article 723.

Straub, P. and Murnighan, K. (1995) An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers. *Journal of Economic Behavior and Organization* 27, 345-364.

Strotz, R.H. (1955) Myopia and inconsistency in dynamic utility maximization. Review of Economic Studies 23, 165-180.

Sundali, J.A., Rapoport, A., and Seale, D.A. (1995) Coordination in market entry games with symmetric players. *Organizational Behavior and Human Decision Processes* 64, 203-218.

Svenson, O. (1981) Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica* 47, 143-148.

Terrell, D. (1994) A test of the gambler's fallacy: evidence from pari-mutuel games. *Journal of Risk and Uncertainty* 8, 309-317.

Tesler, M. (2012) The spillover of racialization into health care: how President Obama polarized public opinion by racial attitudes and race. *American Journal of Political Science* 56(3), 690-704.

Thaler, R. (1980) Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1, 39-60.

Thaler, R.H. (1981) Some empirical evidence on dynamic inconsistency. *Economics Letters* 8, 201-207.

Thaler, R. H. (1985) Mental accounting and consumer choice. *Marketing Science* 4, 199–214.

Thaler, R.H. (2016) Behavioral economics: past, present, and future. *American Economic Review* 106(7), 1577-1600.

Thaler, R.H. and Benartzi, S. (2004) Save More Tomorrow™: using behavioral economics to increase employee saving. *Journal of Political Economy* 112(S1) Papers in Honor of Sherwin Rosen: A Supplement to Volume 112, S164-S187.

Thaler, R.H. and Johnson, E.J. (1990) Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science* 36, 643-660.

Thaler, R.H. and Sunstein, C.R. (2009) *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York, NY: Penguin Books.

Thaler, R.H., Tversky, A., Kahneman, D., and Schwartz, A. (1997) The effect of myopia and loss aversion on risk taking: an experimental test. *Quarterly Journal of Economics* 112, 647-661.

Tiseo, I. (2021) Largest waste producing countries worldwide per capita 2019. Statista. Retrieved from the internet on February 9, 2021 at https://www.statista.com/statistics/1168066/largest-waste-producing-countries-worldwide-per-capita/.

Toussaint, L. and Webb, J.R. (2005) Gender differences in the relationship between empathy and forgiveness. *Journal of Social Psychology* 145(6), 673-685.

Tversky, A. and Kahneman, D. (1986) Rational choice and the framing of decisions. *The Journal of Business* 59(4), Part 2, S251-S278.

Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: a reference dependent model. *The Quarterly Journal of Economics* 106(4), 1039-1061.

Tversky, A. and Kahneman, D. (1992) Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.

Tversky, A. and Shafir, E. (1992) The disjunction effect in choice under uncertainty. *Psychological Science* 3(5), 305-310.

Tversky, A., Slovic, P., and Kahneman, D. (1990) The causes of preference reversal. *American Economic Review* 80, 204-217.

Tworek, H., Beacock, I., and Eseohe, O. (2020) *Democratic Health Communications during Covid-19: A RAPID Response*. Vancouver: UBC Centre for the Study of Democratic Institutions.

United Nations Office on Drugs and Crime (2017) *World Drug Report 2017*. Vienna: United Nations Office on Drugs and Crime.

Van Huyck, J.B., Battalio, R.C., and Beil, R. (1990) Tacit cooperation games, strategic uncertainty, and coordination failure. *American Economic Review* 80, 234-248.

Van Huyck, J.B., Battalio, R.C., and Cook, J. (1997) Adaptive behavior and coordination failure. *Journal of Economic Behavior and Organization* 32, 483-503.

Vanderbuilt, T. (2008) The traffic guru. *The Wilson Quarterly*, Summer, 26-32.

Vincz, T. (2020) Telehealth to tackle addiction by rewarding recovery. *Horizon Health News*. Retrieved from the internet on August 30, 2020 at https://www.horizonhealthnews.com/telehealth-to-tackle-addiction-by-rewarding-recovery/.

Vohs, K.D., Mead, N.L., and Goode, M.R. (2006) The psychological consequences of money. *Science* 314, 1154-1156.

Vuchinich R.E. and Simpson, C.A. (1998) Hyperbolic temporal discounting in social drinkers and problem drinkers. *Experimental Clinical Psychopharmacology* 6, 292-305.

Waber, R.L., Shiv, B., Carmon, Z., and Ariely, D. (2008) Commercial features of placebo and therapeutic efficacy. *Journal of the American Medical Association* 299(9), 1016-1017.

Wald D.S., Bestwick, J.P., Raiman, L., Brendell, R., and Wald, N.J. (2014) Randomised rtial of text messaging on adherence to cardiovascular preventive treatment (INTERACT trial). *PLoS ONE* 9(12), e114268.

Waldfogel, J. (1993) The deadweight loss of Christmas. *American Economic Review* 83(5), 1328-1336.

Wansink, B., Park, S.B., Sonka, S., & Morganosky, M. (2000) How soy labeling influences preference and taste. *International Food and Agribusiness Management Review* 3, 85-94.

Wason, P.C. (1968) Reasoning about a rule. *Quarterly Journal of Experimental Psychology* 20, 273-281.

Weber, R. (2001) Behavior and learning in the "dirty faces" game. *Experimental Economics* 4, 229-242.

Weber, R.A. and Camerer, C.F. (2003) Cultural conflict and merger failure: an experimental approach. *Management Science* 49(4), 400-415.

Wertenbroch, K. (1998) Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science* 17(4), 317-337.

Whittaker, S., Matthews, T., Cerruti, J., Badenes, H., and Tang, J. (2011) Am I wasting my time organizing email?: a study of email refinding. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, May, 3449-3458.

Williams, J. and Wilson, V. (2019) Black workers endure persistent racial disparities in employment outcomes. Economic Policy Institute. Retrieved from the internet on September 30, 2020 at https://www.epi.org/publication/labor-day-2019-racial-disparities-in-employment/.

Wübben M. and von Wangenheim, F. (2008) Instant customer base analysis: managerial heuristics often "get it right." *Journal of Marketing* 72, 82-93.

Yoeli, E., Hoffman, M., Rand, D.G., Nowak, M.A. (2013) Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences (PNAS)* 10(2), 10424-10429.

Zajonc, R. B. and Rajecki, D. W. (1969) Exposure and affect: A field experiment. *Psychonomic Science 17*(4), 216–217.

APPENDIX A - EXAMPLE RESPONSE CARDS

SECTIONS 1 AND 2

Priming Effect

ID Number ______

Did you read about this topic ahead of time? _________

Answer ____________

Anchoring Effect

ID Number ______

Did you read about this topic ahead of time? _________

Question Number _______

Answer ____________________________________

Conjunction Fallacy

ID Number ______

Did you read about this topic ahead of time? _________

Ranking: ____

Testing the Invariance Axiom (version 2) – Experiment 1

ID Number ______

Did you read about this topic ahead of time? _________

Answer ______

Testing the Invariance and Dominance Axioms (version 1)

ID Number ______

Did you read about this topic ahead of time? _________

Experiment 1

Answer ______

Experiment 2

Answer for Compound Lottery 1 ________

Answer for Compound Lottery 2 ________

Fairness in the Context of Framing – Experiment 3

ID Number ______

Did you read about this topic ahead of time? _________

Most Fair Option _________________

Least Fair Option _________________

SECTION 3

Ultimatum Bargaining

ID Number (Proposer) ______

Did you read about this topic ahead of time? _________

ID Number (Responder) ______

Did you read about this topic ahead of time? _________

Value of x ______

Nash Demand Game

ID Number (Player 1) ______

Did you read about this topic ahead of time? _________

ID Number (Player 2) ______

Did you read about this topic ahead of time? _________

Round 1

Number of Aces Dealt to Player 1 ______

Round 2

Did the Players Advance to Round 3? _______

Round 3

Player 1's Demand $ ________

Player 2's Demand $ ________

Finite Alternating-Offer Game

ID Number (Player 1) ______

Did you read about this topic ahead of time? _________

ID Number (Player 2) ______

Did you read about this topic ahead of time? _________

1st Period Offer by Player 1 $ __________

2nd Period Counteroffer by Player 2 $ _________

Final Outcome: Player 1 $ ___________ Player 2 $ __________

Continental Divide Game

Your ID Number ______

Did you read about this topic ahead of time? _________

Group ID Number ______

1st Period Offer __________

2nd Period Counteroffer _________

Final Outcome ___________

Round	My Number	Median	My Payoff
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Escalation Game with Incomplete Information

ID Number (Player 1) ______

Did you read about this topic ahead of time? _________

ID Number (Player 2) ______

Did you read about this topic ahead of time? _________

Player 2's random draw __________

Player 1's 1st period move __________

Player 2’s 2nd period move __________

Police Search

ID Number ______

Did you read about this topic ahead of time? _________

Answer ______

Why? ____

____

Dirty Faces Game

Your ID Number ______

Did you read about this topic ahead of time? _________

Other Player’s ID Number _______

Round	Other Player’s Type	Your Choice	Other Player’s Choice
1			
2			
3			
4			
5			

Stag Hunt

ID Number (Player 1) ______

Did you read about this topic ahead of time? _________

ID Number (Player 2) ______

Did you read about this topic ahead of time? _________

Player 1's Choice __________

Player 2's Choice __________

Trust Game

ID Number (Investor) ______

Did you read about this topic ahead of time? _________

ID Number (Trustee) ______

Did you read about this topic ahead of time? _________

T	$T/2500$	y	$(2T - y)/T$

Centipede Game

ID Number (Player 1) ______

Did you read about this topic ahead of time? _________

ID Number (Player 2) ______

Did you read about this topic ahead of time? _________

	Period					
Player	1	2	3	4	5	6
1		X		X		X
2	X		X		X	

Weakest Link Game

Your ID Number ______

Did you read about this topic ahead of time? _________

Group ID Number ______

	Period									
Results	1	2	3	4	5	6	7	8	9	10
My #										
Low #										
My Pay										

Market Entry Game

Your ID Number ______

Did you read about this topic ahead of time? _________

Group ID Number ______

Your Player Number ___________

	c									
	1	3	5	7	9	11	13	15	17	19
Player										
m										
v										

Weakest Link Game (with Local Interaction)

Your ID Number ______

Did you read about this topic ahead of time? _________

Group ID Number ______

Period										
Results	1	2	3	4	5	6	7	8	9	10
My Choice										
Equilibrium										
My Payoff										

Media Attributions

- Table 1 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 2 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 3 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 4 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 5 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 6 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Table 7 (Appendix A) © Arthur Caplan is licensed under a CC BY (Attribution) license

APPENDIX B - STUDENT SURVEY

Identification Number ________

Please answer the following questions to the best of your ability. If you would prefer not to answer any particular questions, just leave them blank.

(1) What is your gender? (Circle one)

Male   Female

Other (Please explain):

____

(2) Are you employed at least part time? (Circle one)

Yes   No

(3) If you answered "Yes" to Question 2, then describe your work on the lines below, including the number of hours you work per week. If you answered "No," then skip to Question 4.

____

____

(4) What is your current marital status? (Circle one)

Single and never married   Single and divorced   Single and a widow or widower

Married   Cohabitate (live with a partner but not married)

(5) In what year were you born? ___________

(6) How many people currently live in your home, including yourself? ___________

(7) How many children under the age of 5 currently live in your home? ___________

(8) How would you describe your religion and the extent to which you practice it?

____

____

(9) How would you describe your ethnicity? _______________________________

____

(10) What is your current academic status? (Circle only one). If you are a graduate student, circle Graduate Student and skip to Question 11.

Freshman   Sophomore   Junior   Senior   Graduate Student

Other (Please explain):

____

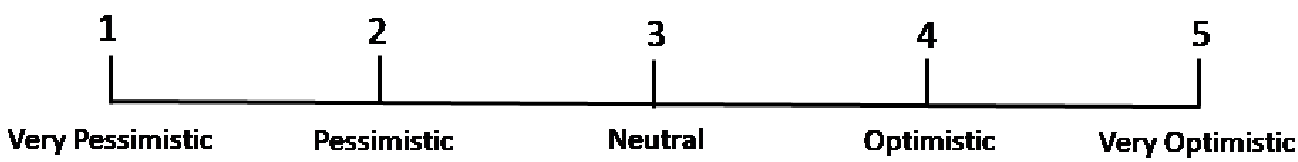
(11) Have you declared a major? _________
If "yes," what is your major? ______________________________
If "no," are you leaning toward choosing a major? ________ If "yes," what major is it?
____

(12) Have you taken any previous economics courses over the course of your entire academic career? _______
If "yes," how many? __________

(13) What is your current cumulative grade point average (GPA)? ____________

(14) What type of employment do you see yourself pursuing after you graduate from the university? You can answer "I don't know" if you are uncertain.
____
____

(15) On a scale from 1 to 5, with 1 corresponding to "Very Pessimistic" and 5 corresponding to "Very Optimistic," how do you feel about the direction your life is currently heading? (Circle one number).



(16) If you could be granted one wish in life, what would it be?
____
____
____

(17) If you were offered the following bet, would you take it?
We flip a fair coin. If it comes up "heads" we pay you $10. If it comes up "tails" you pay us $15.
Yes, I'll take the bet. No, I won't take the bet. Sorry, but I never bet.

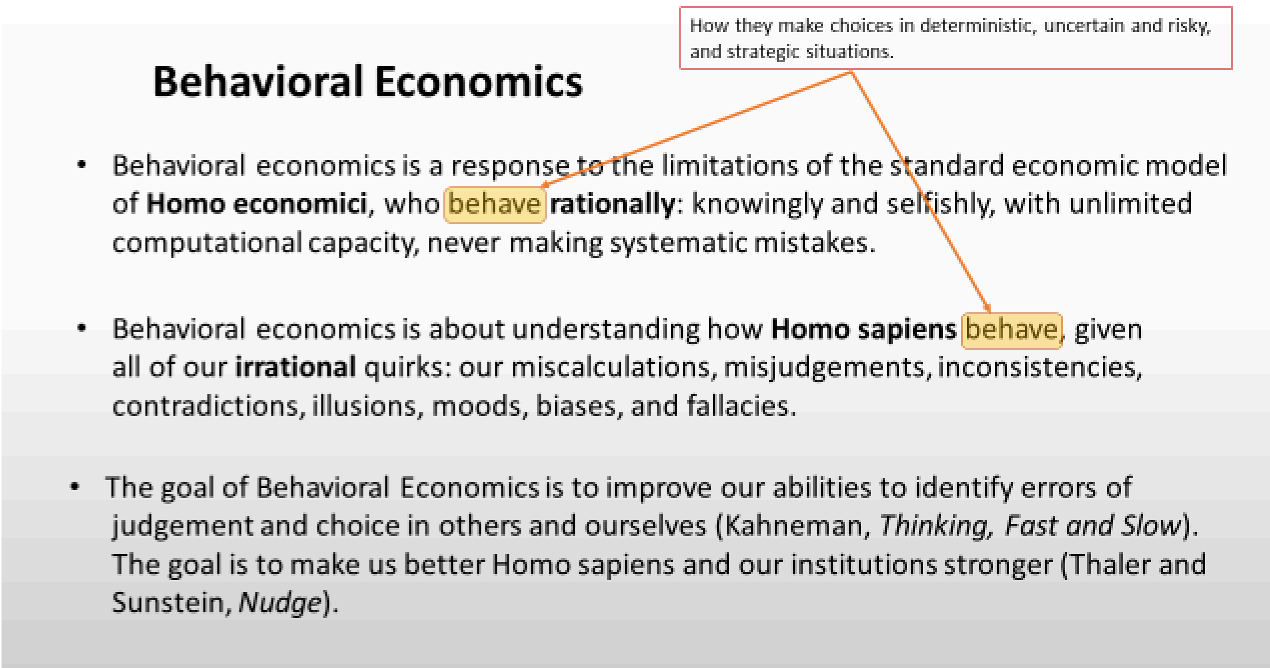
(18) If instead you were offered the following bet, would you take it?
We flip a fair coin. If it comes up "heads" we pay you $10. If it comes up "tails" you pay us $10.
Yes, I'll take the bet. No, I won't take the bet. Sorry, but I never bet.

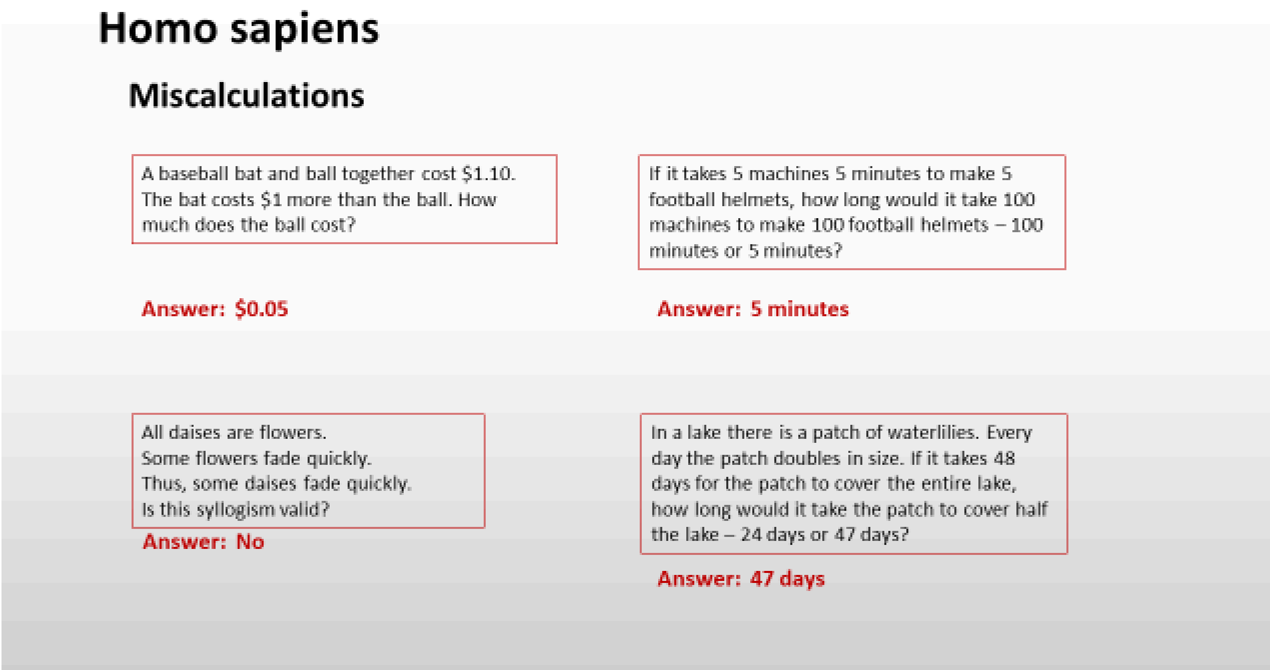
Thank you for participating in this survey!

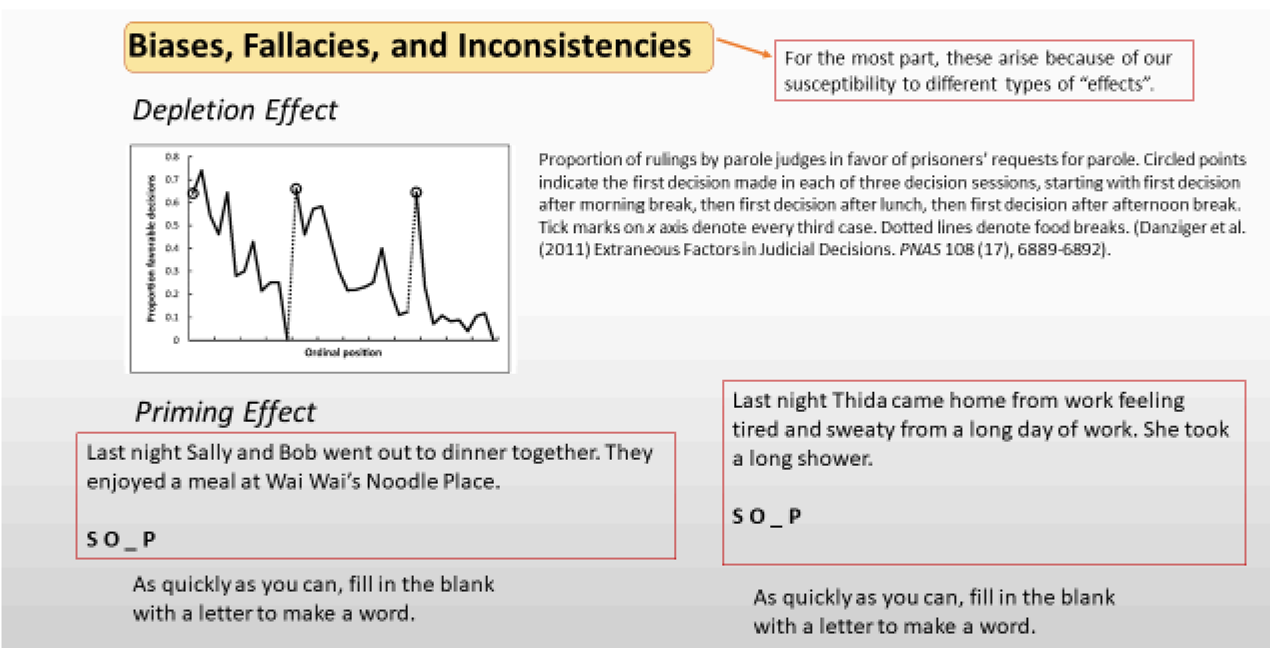
Media Attributions

APPENDIX C - EXAMPLE PRESENTATION SLIDES

Introduction Section







Depletion Effect (Danziger and Levav 2011) Reproduced with permission from publisher.

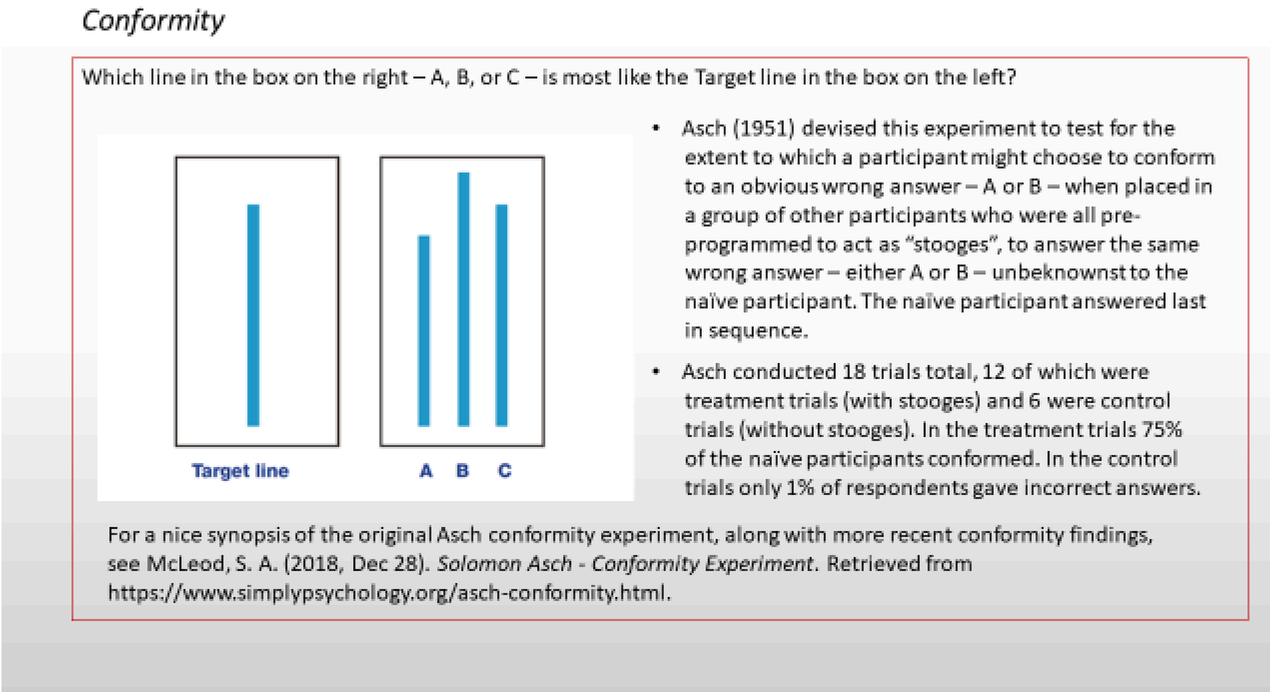
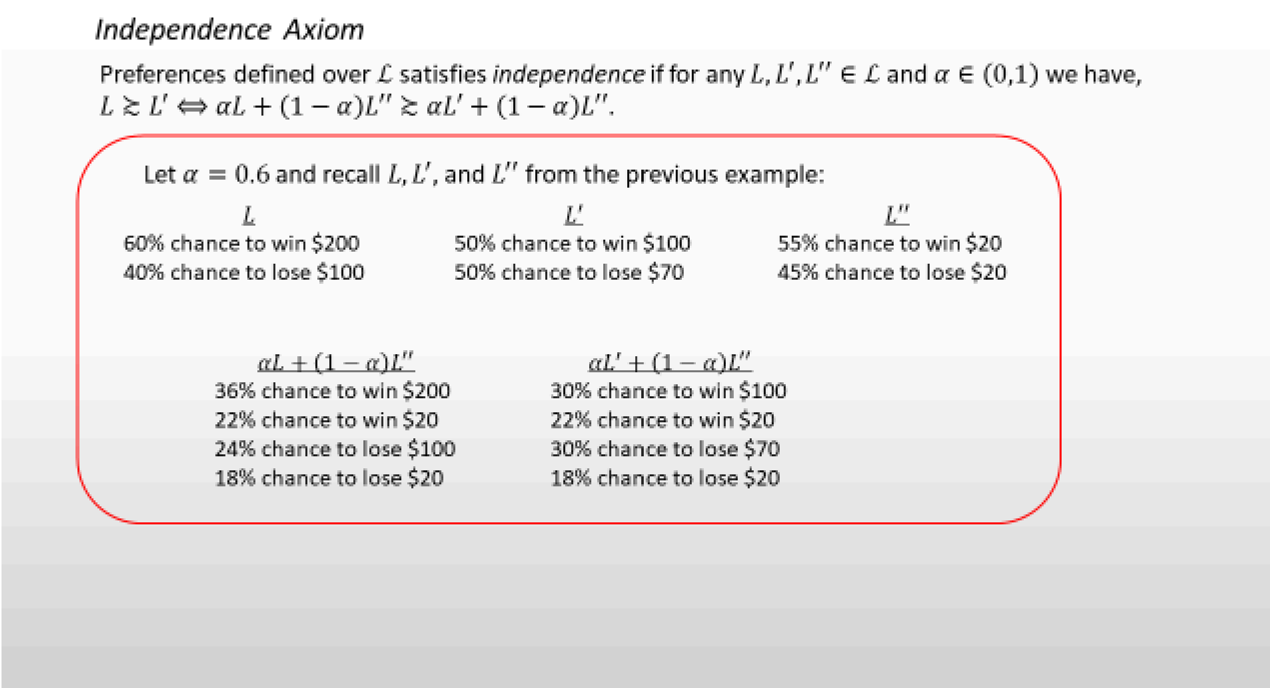


Image: "Psychology-asch-1951", by Saul McLeod, licensed under CC BY 3.0

Chapter 3



Expected Utility Form *

$U: \mathcal{L} \to R$ has *expected utility form* (EUF) over any given lottery if there is an assignment of utility and probability values $u(w_i)$ and p_i, respectively, defined over the lottery's outcomes $i = 1, \ldots, N$, such that $U(L) = \sum_i p_i\, u(w_i)$

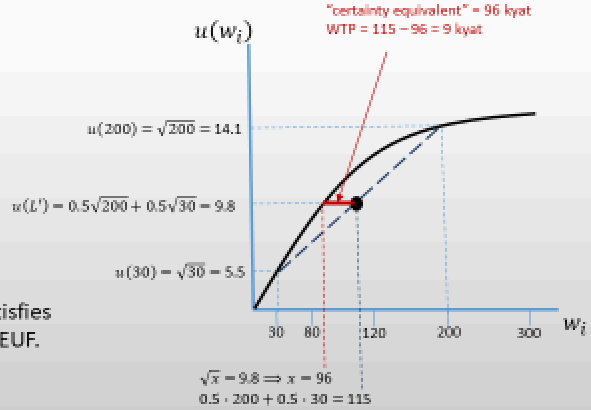
w_i represents "total wealth", the sum of the individual's initial wealth plus(minus) winning(loss) associated with outcome i.

For example, if an individual has initial wealth of 100 kyat and $u(w_i) = \sqrt{w_i}$. Then,

$$U(L) = 0.6\sqrt{300} + 0.4\sqrt{0} = 10.4$$
$$\boxed{U(L') = 0.5\sqrt{200} + 0.5\sqrt{30} = 9.8}$$
$$U(L'') = 0.55\sqrt{120} + 0.45\sqrt{80} = 9.5$$

Thus, $L \succ L' \succ L''$

Graphically,



* It turns out that any utility function $U: \mathcal{L} \to R$ that satisfies the *Independence Axiom* must be represented in the EUF.

Chapter 4

Homo sapiens*

Introduction

Kahneman, Tversky, and Thaler show us why Homo sapiens are,

1. *Risk averse* in the domain of gains, but *loss averse* (and thus *risk seeking*) in the domain of losses.
2. *Reference Dependent* in terms of how we evaluate gains and losses wealth.
3. Likely to *overweight* sure outcomes as well as improbable events, relative to events of moderate probability.
4. Susceptible to *framing*, where decision problems described in different ways give rise to different preferences (contrary to the invariance axiom governing Homo economicus' choices).
5. Susceptible to *emotions* such as envy and guilt.
6. *Mental accountants*, which explains why the acceptability of a lottery can depend upon whether a negative outcome is evaluated as a cost or as an uncompensated loss.

* These, and the ensuing findings for Homo sapiens, are taken primarily from:
Kahneman, D. and A. Tversky (1984) "Choices, Values, and Frames". *American Psychologist* 39(4), 341-350.
Kahneman, D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision Under Risk". *Econometrica* 47(2), 263-292.
Tversky, A. and D. Kahneman (1986) "Rational Choice and the Framing of Decisions". *The Journal of Business* 59(4), part 2, S251-S278.
Tversky, A. and E. Shafir (1992) "The Disjunction Effect in Choice Under Uncertainty." Psychological Science 3(5), 305-309.
Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics* 75, 643-669.
Heath, C. and A. Tversky (1991) "Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty." *Journal of Risk and Uncertainty* 4(1), 5-28.
Grether, D.M. and C.R. Plott (1979) "Economic Theory of Choice and the Preference Reversal Phenomenon." *American Economic Review* 69(4), 623-638.

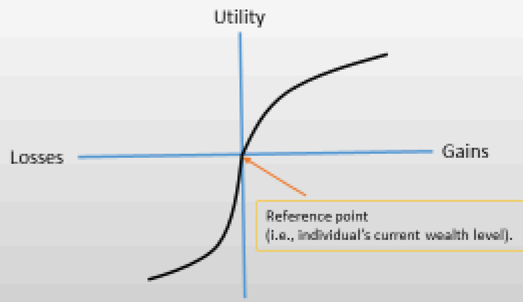
The Value Function

People do not normally consider relatively small outcomes in terms of total wealth, but rather in terms of gains, losses, and neutral outcomes, e.g., maintenance of the status quo.

And just as utility value is a concave function of the size of a gain, the same can be said of losses, i.e., the difference in utility value between a loss of $200 and a loss of $100 appears greater than the utility value between a loss of $1,200 and a loss of $1,100. This is known as *diminishing sensitivity* in the realm of losses.

To the extent that people suffer from *loss aversion*, the concave function defined over losses is steeper than that defined over gains (i.e., a loss of $X is more averse than a gain of $X is attractive).

The Value Function



Chapter 5

Violation of Invariance and Dominance

Experiment 1

The following lottery is described by the percentage of marbles of different colors in each box and the amount of money you win or lose depending upon the color of a randomly drawn marble. Which lottery do you prefer?

A	90% white $0	6% red $45,000	1% green $30,000	1% blue -$15,000	2% yellow -$15,000
B	90% white $0	6% red $45,000	1% green $45,000	1% blue -$10,000	2% yellow -$15,000

Answer? **0% in favor of A, 100% in favor of B** (N = 88)

Experiment 2

The following lottery is described by the percentage of marbles of different colors in each box and the amount of money you win or lose depending upon the color of a randomly drawn marble. Which lottery do you prefer?

C	90% white $0	6% red $45,000	1% green $30,000	3% yellow -$15,000
D	90% white $0	7% red $45,000	1% green -$10,000	2% yellow -$15,000

Answer? **58% in favor of C, 42% in favor of D** (N = 124)

Notes:

1. In Experiment 1, lottery **B** dominates lottery **A**.
2. Experiment 2 is effectively identical to Experiment 1, except that colors yielding identical outcomes (red and green marbles in lottery **B** and yellow and blue in lottery **A**) are combined, i.e., lottery **C** is identical to **A** and lottery **D** is identical to **B**.

Key Result: The outcome of Tversky and Kahneman's (1986) experiment was..... No violation of *Dominance* in Experiment 1 alone, since all subjects chose option **B** over option **A**. But in Experiment 2 a slight majority of subjects chose option **C** over option **D** which, as pointed out in Note 2, implies a preference reversal. This is a violation of both the *Dominance and Invariance Axioms*!

Violation of Sure-Thing Principle

Experiment 7a

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and run-down, and you are not sure that you passed the exam. If you failed you will have to take the exam again in a couple of months – after the semester break. You now have an opportunity to buy a very attractive 5-day vacation package to Goa Beach in India at an exceptionally low rice. The special offer expires tomorrow, while the exam grade will not be available until the day after tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a 50,000 kyat non-refundable fee in order to retain the right to buy the vacation package at the same low price the day after tomorrow – after you learn whether you passed the exam.

Answer? **32% chose A, 7% chose B, 61% chose C** (N = 66)

Experiment 7b

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and run-down, and you find out that you passed the exam. You now have an opportunity to buy a very attractive 5-day vacation package to Goa Beach in India at an exceptionally low rice. The special offer expires tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a 50,000 kyat non-refundable fee in order to retain the right to buy the vacation package at the same low price the day after tomorrow – after you learn whether you passed the exam.

Answer? **54% chose A, 16% chose B, 30% chose C** (N = 67)

Experiment 7c

Imagine you have just taken a difficult examination. It is the end of the fall semester, you feel tired and run-down, and you find out that you failed the exam. You will have to take the exam again in a couple of months – after the semester break. You now have an opportunity to buy a very attractive 5-day vacation package to Goa Beach in India at an exceptionally low rice. The special offer expires tomorrow. Would you:

A Buy the vacation package.
B Not buy the vacation package.
C Pay a 50,000 kyat non-refundable fee in order to retain the right to buy the vacation package at the same low price the day after tomorrow – after you learn whether you passed the exam.

Answer? **57% chose A, 12% chose B, 31% chose C** (N = 67)

Note: If the percentages of those choosing **A**, **B**, and **C** in Experiment 7b are roughly equal to their corresponding percentages in Experiment 7c, then these same percentages should be roughly equal those in Experiment 7a.

Key Result: The outcome of Tversky and Shafir (1992) experiment was….. an apparent violation of the *Sure-Thing Principle*!

Chapter 6

Overweighting Improbable Events

Experiment 11a

Choose between lotteries **A** and **B**:

A 0.001% chance to win 50,000 kyat
B win 5,000 kyat for certain

Answer? **72% A, 28% B** (N = 72)

Suggests why we are prone to gamble.

Experiment 11b

Choose between lotteries **A** and **B**:

A 0.001% chance to lose 50,000 kyat
B lose 5,000 kyat for certain

Answer? **17% A, 83% B** (N = 72)

Suggests why we purchase insurance policies.

Vividness of Probability

Experiment 1

Suppose you are a Psychiatrist at a psychiatric hospital. You are in charge of evaluating whether it is safe to discharge Mr. Thiha from the hospital. Mr. Thiha has a history of violence. You have received the following assessment from a criminal expert concerning the risk associated with releasing Mr. Thiha from the hospital:

> "Patients similar to Mr. Thiha are estimated to have a 10% probability of committing an act of violence against others during the first several months after discharge."

Will you deny Mr. Thiha's discharge?

Answer? **21% Yes 79% No**

Experiment 2

Suppose you are a Psychiatrist at a psychiatric hospital. You are in charge of evaluating whether it is safe to discharge Mr. Thiha from the hospital. Mr. Thiha has a history of violence. You have received the following assessment from a criminal expert concerning the risk associated with releasing Mr. Thiha from the hospital:

> "Of every 100 patients similar to Mr. Thiha, 10 are estimated to commit an act of violence against others during the first several months after discharge."

Will you deny Mr. Thiha's discharge?

Answer? **41% Yes 59% No**

Key Result: The outcome of Kahneman's experiment (reported in *Thinking, Fast and Slow*) was.....which is an example of a *framing effect*, since the percentages should be equal across Experiments 1 and 2.

Testing for an Endowment Effect

Experiment 17

10 of you have been given a [To Be Determined]. You are henceforth known as "Sellers".
10 of you have not been given these items. You are henceforth known as "Buyers".

Each of the Sellers will now write on their piece of paper their "Seller's Price"; the price at which they would willingly sell their [To Be Determined] to one of the Buyers.
Each of the Buyers will now write on their piece of paper their "Buyer's Price"; the price at which they would willingly buy [To Be Determined] from one of the Sellers.

Seller's prices will be ranked from highest to lowest and then compared with the highest-to-lowest ranking of Buyer's prices to determine which trades will occur via paired bids.

Example of How Trades Are Determined

Assume the good being traded among five Sellers and five Buyers is a beautiful coffee cup.
The rankings of Sellers and Buyers Prices are,

Seller/ Buyer #	Sellers Prices ($)	Buyers Prices ($)
1	2.00	2.20
2	1.50	1.50
3	1.30	1.20
4	1.00	0.50
5	0.50	0.50

Here, Buyer 1 would pay Seller 1's price of $2.00.
Buyer 2 would pay Seller 2's price of $1.50.
Buyer 3 would pay Seller 4's price of $1.00.
Buyer 4 would pay Seller 5's price of $0.50.

Our Results →

Seller/ Buyer #	Sellers Prices	Buyers Prices
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Ultimatum Bargaining

Two players – a Proposer and a Responder – bargain over $100. The Proposer offers some portion, x, of the $100 to the Responder, leaving the Proposer with $\$(100 - x)$. If the Responder accepts the offer, then she gets $\$x$ and the Proposer gets $\$(100 - x)$. If the Responder rejects the offer, both players get nothing.

Outcome?

Homo economicus: By going first, the Proposer has all of the bargaining power. The Proposer therefore exploits the fact that the self-interested Responder will take whatever is offered. The amount offered is thus very close to zero.

Homo sapiens: In a multitude of experiments conducted all over the world, Proposers typically offer roughly half of the total. Offers of roughly 20% are rejected about half of the time as punishment for not behaving fairly ("negative reciprocity"). If there is more than one Proposer equilibrium, $x \rightarrow$ total amount.

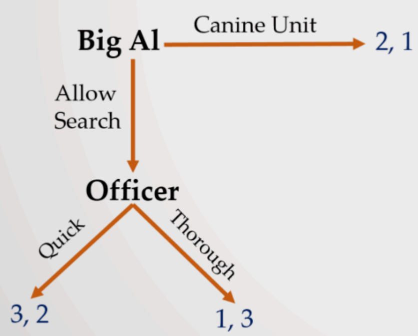
Ironically, primitive cultures in Africa, the Amazon, Papua New Guinea, Indonesia, and Mongolia behave more like Homo economicus than Homo sapiens!

Police Search

Suppose a police officer pulls Big Al over and asks to search his vehicle. Big Al can let the police officer search the vehicle (which could be a quick or thorough search, depending upon the police officer's preferences) or refuse and force the officer to call in the Canine Unit. Big Al's preferences are: Quick Search ≻ Canine Unit ≻ Thorough Search, while the police officer's preferences are: Thorough Search ≻ Quick Search ≻ Canine Unit.

The officer points out to Big Al that a "Quick Search is more preferred for both us than calling in the Canine Unit".

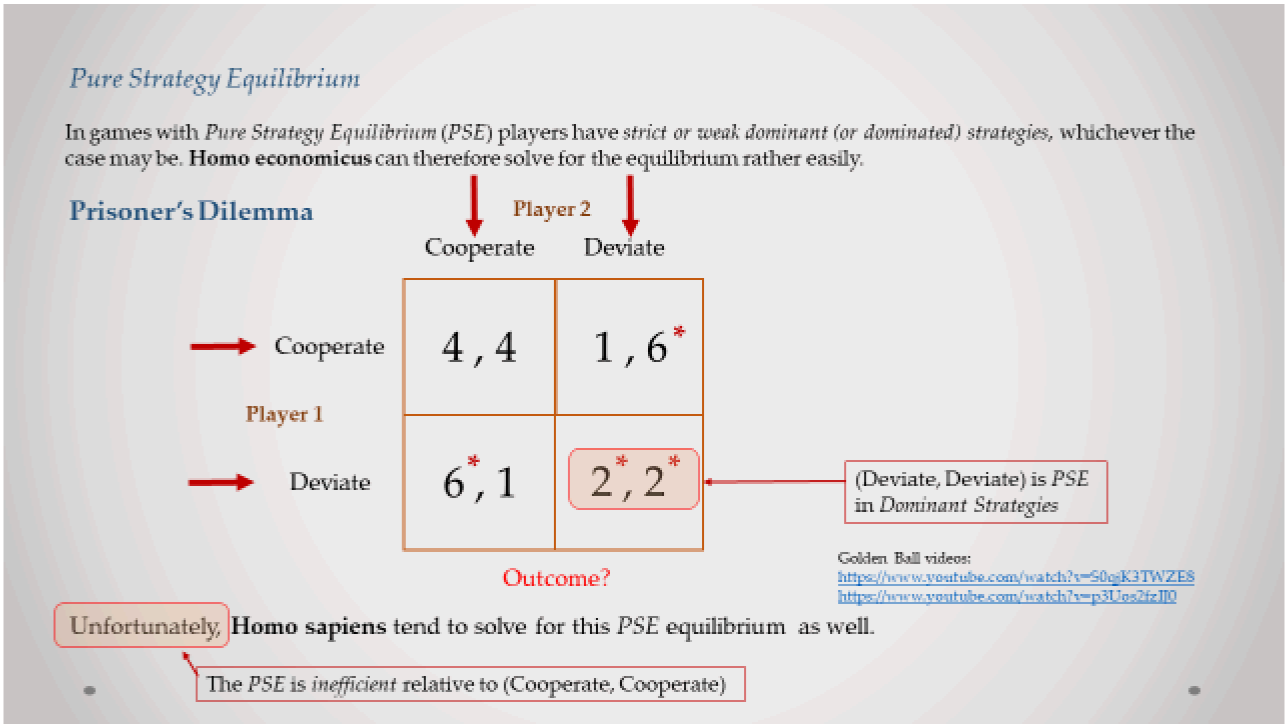
Should Big Al agree to let the police officer search his vehicle? **No**

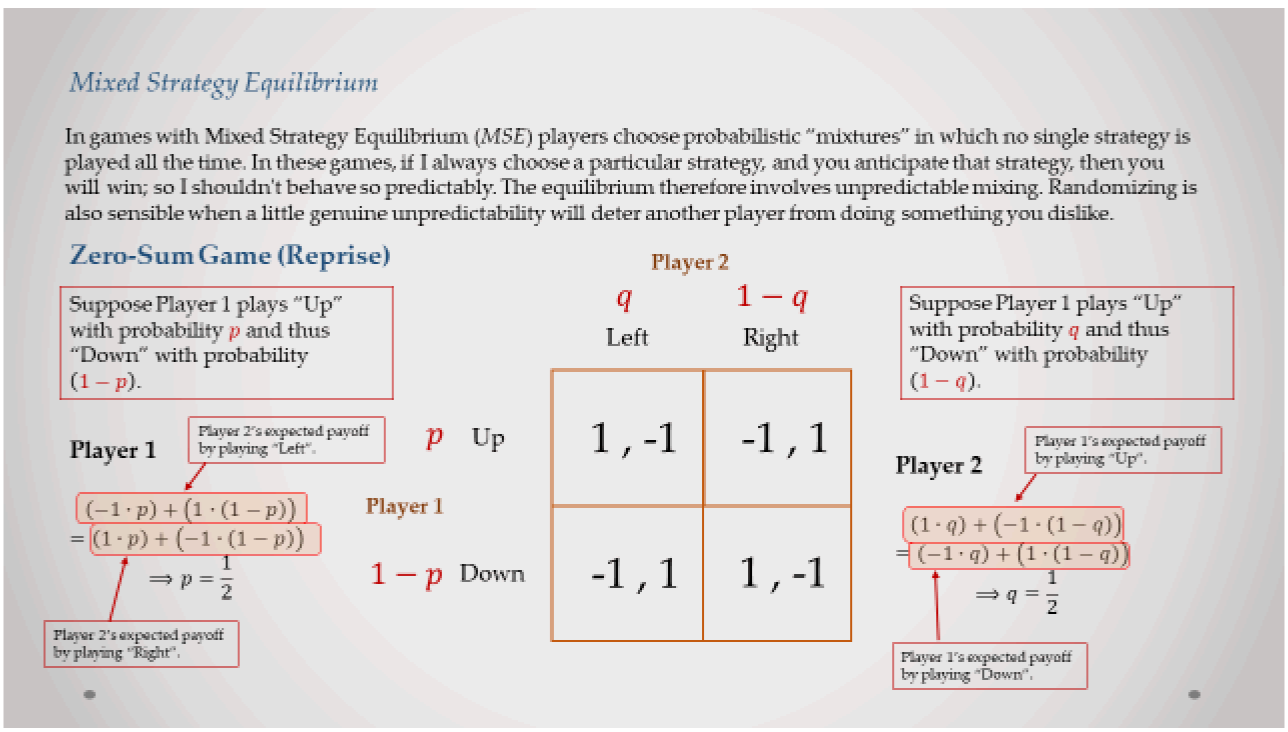


If Big Al allows a search, the officer will choose "Thorough", implying Big Al's payoff is **1** and the officer's is **3**.

If Big Al instead does not allow the search, the Canine Unit is called in, implying Big Al's payoff is **2** and the officer's is **1**.

Thus, this game's *SPE* is Big Al not allowing a search and the officer calling in the Canine Unit.


Pure Strategy Equilibrium
In games with Pure Strategy Equilibrium (PSE) players have strict or weak dominant (or dominated) strategies, whichever the case may be. Homo economicus can therefore solve for the equilibrium rather easily.
Prisoner's Dilemma
Player 2
Cooperate
Deviate
Cooperate
Player 1
Deviate
4 , 4
1 , 6*
6*, 1
2*, 2*
(Deviate, Deviate) is PSE in Dominant Strategies
Outcome?
Golden Ball videos:
https://www.youtube.com/watch?v=S0qjK3TWZE8
https://www.youtube.com/watch?v=p3Uos2fzIl0
Unfortunately, Homo sapiens tend to solve for this PSE equilibrium as well.
The PSE is inefficient relative to (Cooperate, Cooperate)


Mixed Strategy Equilibrium
In games with Mixed Strategy Equilibrium (MSE) players choose probabilistic "mixtures" in which no single strategy is played all the time. In these games, if I always choose a particular strategy, and you anticipate that strategy, then you will win; so I shouldn't behave so predictably. The equilibrium therefore involves unpredictable mixing. Randomizing is also sensible when a little genuine unpredictability will deter another player from doing something you dislike.
Zero-Sum Game (Reprise)
Suppose Player 1 plays "Up" with probability p and thus "Down" with probability (1 − p).
Player 2
q
1 − q
Left
Right
Suppose Player 1 plays "Up" with probability q and thus "Down" with probability (1 − q).
Player 1
Player 2's expected payoff by playing "Left".
p Up
1 , -1
-1 , 1
Player 2
Player 1's expected payoff by playing "Up".
(−1 · p) + (1 · (1 − p))
= (1 · p) + (−1 · (1 − p))
⇒ p = 1/2
Player 1
1 − p Down
-1 , 1
1 , -1
(1 · q) + (−1 · (1 − q))
= (−1 · q) + (1 · (1 − q))
⇒ q = 1/2
Player 2's expected payoff by playing "Right".
Player 1's expected payoff by playing "Down".

Discrimination on the Weakest Link Game Show? *

- Contestant voting behavior on the television game show **Weakest Link** provides an unusual opportunity for the author to distinguish between *taste-based* (bad!) and *information-based* (not-as-bad!) theories of discrimination.
- Taste-based discrimination occurs when people prefer not to interact with a particular class of people and are willing to pay a financial price to avoid such interactions. With information-based discrimination people have no animus against a particular class of people (unlike taste-based models), but discriminatory outcomes nonetheless arise because they have less reliable (i.e., noisy) information about them.
- Let's watch these two videos to learn how the Weakest Link game show is actually played. Video 1 and Video 2.
  Note that contestants answer trivia questions over a series of rounds, with one contestant eliminated each round on the basis of the votes of the other contestants, until only two contestants remain. The last two contestants compete head-to-head for the winner-take-all prize. Because the prize money at stake is potentially large (the money is an increasing function of the number of questions answered correctly over the course of the program) participants have powerful incentives to vote in a manner that maximizes their chance of winning.
- In early rounds, strategic incentives encourage voting for the weakest competitors. In later rounds, the incentives reverse and the strongest competitors become the logical target. Both theories of discrimination predict that in early rounds excess votes will be cast against people targeted for discrimination. In later rounds, however, taste-based models predict continued excess votes, whereas statistical discrimination predicts fewer votes against the targeted people.
- Players do vote strategically in early rounds, but not in middle and final rounds.
  Evidence of discrimination is limited. There is little in the data to suggest discrimination against women and African Americans. Some evidence suggests information-based discrimination toward Hispanics and taste-based discrimination against older players.

Regression Analysis of Votes Received

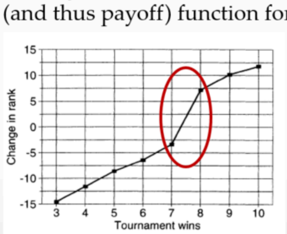
Variable	Early Rounds (N = 1,191) (1)	(2)	Middle Rounds (N = 1,599) (3)	(4)	Final Round (N = 483) (5)	(6)
Female	−.09 (.09)	−.07 (.08)	.00 (.06)	−.04 (.05)	−.06 (.08)	−.10 (.08)
Black	.13 (.11)	.07 (.09)	−.05 (.08)	−.00 (.06)	−.01 (.09)	−.01 (.09)
Asian	−.34 (.17)	−.20 (.13)	.28 (.19)	.33 (.15)	−.17 (.24)	−.13 (.26)
Hispanic	.32 (.29)	.17 (.24)	.40 (.25)	.41 (.19)	−.34 (.28)	−.35 (.24)
Age 50+	.34 (.19)	.17 (.17)	.30 (.13)	.30 (.11)	.49 (.16)	.42 (.16)
% Correct this round (deviation from other players)	. . .	−2.44 (.11)	. . .	−1.71 (.09)	. . .	−.47 (.16)
Squared % correct this round (deviation from other players)	. . .	2.53 (.30)	. . .	.73 (.25)	. . .	.24 (.37)
Cumulative % correct (deviation from other players)	. . .	−.90 (.38)	. . .	−.70 (.14)	. . .	−.33 (.33)
Squared cumulative % correct (deviation from other players)	. . .	1.37 (.88)	. . .	.59 (.33)	. . .	2.77 (1.47)
Cumulative opponents voted against who are still alive	. . .	.12 (.29)	. . .	.27 (.04)	. . .	.09 (.04)
R^2	.013	.387	.010	.257	.043	.089
Education and region dummies included?	No	Yes	No	Yes	No	Yes

Note.—The dependent variable in all columns is the number of votes received by a contestant in a given round. The unit of observation is a contestant round. Estimation is done with ordinary least squares. Standard errors, clustered by episode and round, are in parentheses. All regressions include an exhaustive set of interactions controlling for round by show-length interactions. The even columns include state-of-residence fixed effects and occupation dummies, but these coefficients are not reported in the table. Early rounds correspond to the first two rounds of the prime-time show (1 hour long) and the first round of the daytime show (30 minutes long). Middle rounds correspond to rounds 3–5 of the -time show and rounds 2 and 3 of the daytime show. Final round corresponds to round 6 of the prime-time show and round 4 of the daytime show.

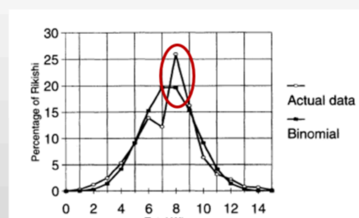
* Levitt, S.D. (2004) "Testing Theories of Discrimination: Evidence from Weakest Link". *Journal of Law and Economics* 48, 431-452.

Corruption in Sumo Wrestling *

- A sumo tournament involves 66 wrestlers (*rikishi*) competing in 15 bouts each. A wrestler who achieves a winning record (eight wins or more) is guaranteed to rise in the official ranking. A wrestler with a losing record (seven wins or less) falls in the rankings. A wrestler's ranking is a source of prestige, and the basis for salary determination and perks.
- The key institutional feature of sumo wrestling that makes it ripe for corruption is the sharp nonlinearity in the ranking (and thus payoff) function for competitors:



- Consequentially, a wrestler entering the final match of a tournament with a 7-7 record has far more to gain from a victory than an opponent with a record of, say, 8-6 has to lose.
- Following almost 300 wrestlers from 1989-2000, the authors find that wrestlers who are on the margin for attaining their eighth victory in a tournament (in what's known as a "bubble match") win far more often than would be expected:
- Whereas the wrestler who is on the margin for an eighth win is victorious with a surprisingly high frequency, the next time the same two wrestlers face each other, it is the opponent who has an unusually high win percentage.



- The bubble match effect disappears in tournaments with a high level of media scrutiny and when the opponent is in the running for a special prize.
- Success on the bubble increases for veterans.

* Duggan, M. and S.D. Levitt (2002) "Winning Isn't Everything: Corruption in Sumo Wrestling". *American Economic Review* 92(5), 1594-1605.

Media Attributions

- Slide 1 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 2 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 3 (Appendix C) © Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso; 2021 National Academy of Sciences adapted by Arthur Caplan is licensed under a All Rights

Reserved license

- Slide 4 (Appendix C) © Saul McLeod adapted by Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 5 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 6 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 7 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 8 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 10 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 10a (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 9 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 12 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 14 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 15 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 16a (Appendix C)
- Slide 17 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 18 (Appendix C) © Arthur Caplan is licensed under a CC BY (Attribution) license
- Slide 19a (Appendix C)
- Slide 20a (Appendix C)

APPENDIX D - EXAMPLE COURSE OUTLINES

Course for Economics Majors

Week	Meeting	Topics
1	1	Introduction to Course Assignment of Student ID Numbers Completion of Demographic Surveys
	2	Miscalculations (Chapter 1) Cognitive Illusions (Chapter 1) Heuristics (Chapter 1) Effects (Chapter 1)
2	1	Effects (Chapter 1) Biases and Fallacies (Chapter 2)
	2	Biases and Fallacies (Chapter 2)
3	1	**Short Quiz over Chapters 1 and 2** Principal and Additional Rationality Axioms (Chapter 3)
	2	Principal and Additional Rationality Axioms (Chapter 3) *Homo economicus* and the Expected Utility Form (Chapter 3)
4	1	*Homo economicus* and the Expected Utility Form (Chapter 3) *Homo economicus* and the Indifference Curve (Chapter 3)
	2	*Homo economicus* and the Indifference Curve (Chapter 3) Key Takeaways on *Homo economicus* (Chapter 3)
5	1	Prospect Theory (Chapter 4)
	2	Prospect Theory (Chapter 4) Key Takeaways on *Homo sapiens* (Chapter 4)

6	1	**Exam 1 over Chapters 3 and 4**
	2	Laboratory Experiments: The Rationality of *Homo economicus* Versus the Reality of *Homo sapiens* (Chapter 5)
7	1	Laboratory Experiments: Additional Differences Between *Homo economicus* and *Homo sapiens* (Chapter 6)
	2	Laboratory Experiments: Additional Differences Between *Homo economicus* and *Homo sapiens* (Chapter 6)
8	1	Some Classic Games of Iterated Dominance (Chapter 7)
	2	Some Classic Games of Iterated Dominance (Chapter 7)
9	1	Some Classic Games of Iterated Dominance (Chapter 7)
	2	Some Classic Games of Iterated Dominance (Chapter 7)
10	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)
11	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)
12	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)
13	1	**Exam 2 over Chapters 5 – 8**
	2	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
14	1	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
	2	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
15	1	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
	2	**Short Quiz over Section 4**

Course for Non-Economics Majors

Week	Meeting	Topics
1	1	Introduction to Course Assignment of Student ID Numbers Completion of Demographic Surveys
	2	Miscalculations (Chapter 1) Cognitive Illusions (Chapter 1) Heuristics (Chapter 1)
2	1	Heuristics (Chapter 1) Effects (Chapter 1)
	2	Effects (Chapter 1)
3	1	Biases and Fallacies (Chapter 2)
	2	Biases and Fallacies (Chapter 2)
4	1	**Exam 1 over Chapters 1 and 2**
	2	*Homo economicus* and the Expected Utility Form (Chapter 3)
5	1	*Homo economicus* and the Indifference Curve (Chapter 3)
	2	Prospect Theory (Chapter 4)
6	1	Laboratory Experiments: The Rationality of *Homo economicus* Versus the Reality of *Homo sapiens* (Chapter 5)
	2	Laboratory Experiments: Additional Differences Between *Homo economicus* and *Homo sapiens* (Chapter 6)
7	1	Laboratory Experiments: Additional Differences Between *Homo economicus* and *Homo sapiens* (Chapter 6)
	2	**Exam 2 over Chapters 3 – 6**
8	1	Some Classic Games of Iterated Dominance (Chapter 7)
	2	Some Classic Games of Iterated Dominance (Chapter 7)
9	1	Some Classic Games of Iterated Dominance (Chapter 7)
	2	Some Classic Games of Iterated Dominance (Chapter 7)
10	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)
11	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)
12	1	Some Classic Simultaneous-Move Games (Chapter 8)
	2	Some Classic Simultaneous-Move Games (Chapter 8)

13	1	**Exam 2 over Chapters 7 and 8**
	2	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
14	1	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
	2	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
15	1	Empirical Studies, Field Experiments, and Choice Architecture (Section 4)
	2	**Short Quiz over Section 4**

Media Attributions

APPENDIX E - LINKAGES MATRIX

Linkages Matrix (This link downloads the matrix in an excel file)