

Learning Logic: Examining the Effects of
Context Ordering on Reasoning About Conditionals

Christina W. Lommatsch¹

Christina.Lommatsch@canberra.edu.au.com

Patricia S. Moyer-Packenham²

Patricia.Moyer-Packenham@usu.edu

¹University of Canberra, corresponding author

²Utah State University

Abstract

Logical statements are prevalent in mathematics, the sciences, law, and many areas of everyday life. The most common logical statements are conditionals, which have the form “If H..., then C...,” where “H” is a hypothesis (or condition) to be satisfied and “C” is a conclusion to follow. Reasoning about conditionals depends on four main conditional contexts (e.g., intuitive, abstract, symbolic, or counterintuitive). The purpose of this study was to test a theory about the effects of context ordering on reasoning about conditionals. Researchers developed and tested a virtual manipulative mathematics app, called the Learning Logic App.

This study addressed the question: “How does the order of teaching four conditional contexts influence reasoning about conditionals?” A total of 154 participants, randomly assigned to a context ordering, interacted with the Learning Logic App online. Researchers collected data using a Conditional Logic Assessment, score logs, and participant surveys. The results suggest that context ordering does influence learners’ reasoning. The most beneficial context ordering for learners’ performance was symbolic-intuitive-abstract-counterintuitive. The most beneficial context ordering according to learners’ perceptions was intuitive-abstract-counterintuitive-symbolic. Based on these results, researchers propose that the symbolic-intuitive-abstract-counterintuitive-symbolic ordering may be most effective for learning about conditionals. This progression incorporates a catalyst at the beginning (symbolic) which aids the learner in reassessing prior knowledge. Next, the ordering progresses from easiest to hardest (intuitive-abstract-counterintuitive-symbolic). These findings suggest the importance of the sequencing for teaching and learning to reason about conditionals.

Learning Logic: Examining the Effects of Context Ordering on Reasoning About Conditionals

Logical inference abilities are essential to deep mathematical understanding [1] and can be applied in almost every academic discipline from theorems in chemistry, physics, and biology to programming in computer science. Logical inference abilities have their foundation in reasoning about conditionals. Conditionals are statements of the form “if H, then C,” where H is the hypothesis (or condition) and C is the conclusion. The prevalence of conditionals in academics and in everyday life supports the need for individuals to understand their underlying logical structures.

Reasoning about conditionals can be a difficult task and becomes more demanding depending on the structures of the hypothesis and conclusion and the type of reasoning involved [2]. Reasoning about conditionals rarely develops spontaneously [3] and differs depending on the context of the conditional. Different contexts include those that are intuitive, abstract, symbolic, or counterintuitive. These factors can lead to several types of fallacious reasoning, such as *child’s logic* [4] where an individual assumes the bidirectionality of a conditional. For example, the transformation of the intuitive context conditional “If it is a fish, then it swims” into two conditionals: “If it is a fish, then it swims” and “If it swims, then it is a fish,” would imply that everything that can swim is a fish. This illogical leap can be harmless and somewhat entertaining as in the example above but becomes a much larger concern when it appears in an individual’s interpretation of tax law or ballot measures on which she or he intends to vote. Thus, it is essential for individuals to be able to correctly reason about conditionals.

In addition to the benefits of understanding conditionals for their direct applications,

research has shown that there are correlations of correct conditional reasoning and higher-level reasoning in other fields. For example, Kılıç and Sağlam [5] demonstrated that individuals who could correctly reason about conditionals had higher abilities in comprehension of genetics concepts than their peers who were unable to reason correctly about conditionals. Kılıç and Sağlam further noted that individuals who were unable to correctly reason about conditionals had a much higher tendency to use rote memorization techniques rather than a conceptual understanding approach to learning the science material. Consequently, learning to reason about conditionals may be beneficial not only in direct applications, but also in positively influencing an individual's learning strategies in other areas.

Research shows that reasoning about conditionals is highly context-dependent (e.g., [6-7]). Additionally, there appears to be a progression of difficulty from one context to another (e.g., [8-9]), and that learners do not naturally progress from reasoning correctly in one context to another (e.g., [3, 10]). This is important because this type of reasoning is foundational to (e.g., [11]) and correlated with success in other subject areas [5]. Given the importance of reasoning about conditionals, the purpose of this study was to test a theory about the effects of context ordering on reasoning about conditionals.

Review of the Literature

Historically, the research on reasoning about conditionals has fallen into four main contexts, categorized as abstract, intuitive, symbolic, and counterintuitive. Recent research has focused on comparing individuals' reasoning in each of these four contexts and how learners transition among the contexts.

Four Conditional Contexts

Abstract context conditionals are defined as conditionals where a learner has no prior knowledge that may influence their ability to correctly reason about the conditional [12]. The most well-known study related to reasoning about conditionals is the Wason Selection Task, proposed and administered by P. C. Wason in 1966. The Selection Task asks participants to consider four envelopes, laid flat, that have a letter on one side and a number on the other. The researcher asked participants to “select just the envelopes that definitely needed to be turned over to find out whether they violate the rule” from the four envelopes labeled “D,” “C,” “5,” and “4” [12]. The rule, that is the conditional, proposed was “If an envelope has a D on one side, then it has a 5 on the other side” [12]. Wason found that less than 10% of the undergraduate participants correctly chose both D (the correct solution for direct reasoning) and 4 (the correct solution for contrapositive reasoning). Nearly all participants correctly selected D, however, it was the second card, 4, which would validate the contrapositive of the statement, that most participants failed to select. The Selection Task requires participants to reason about two different logical structures, direct reasoning and contrapositive reasoning, with each of these placed in an abstract context. This means that a participant would have no prior knowledge which could influence their reasoning about the conditional. Studying participants’ responses to this abstract conditional clearly showed that participants had not naturally developed some aspects of reasoning about the conditionals. However, Wagner-Egger [13] found that, when college students were presented with a question equivalent to the Wason Selection Task, that was situated in a different context (intuitive) where they could use prior knowledge to aid their contrapositive reasoning, participants performed significantly better than they had on the Wason Selection Task. This demonstrates that, while the Wason Selection Task may measure a participant’s ability to reason

about a conditional in an abstract context, it fails to generalize to a participants' ability to reason about a conditional in other contexts, such as the intuitive context tested by Wagner-Egger.

Intuitive context conditionals are defined as conditionals where a learner may have prior knowledge that positively influences their ability to correctly reason about the conditional, such as a well-known social rule [6]. Cosmides and Tooby [6] changed the context of the abstract task posed by Wason to a more intuitive context. That is, a participant could use their prior knowledge of a conditional such as a law and the associated social situations to aid their direct and contrapositive reasoning. Cosmides and Tooby concluded that questions involving these social contexts positively influenced individuals' ability to correctly reason about conditionals.

Symbolic context conditionals are defined as conditionals where the learner must reason with mathematics expressions which have no mathematical meaning. Understanding the underlying logical structures of conditionals appears to be a key aspect of transferring knowledge of reasoning about conditionals to contexts other than the intuitive ones studied by Cosmides and Tooby [6] and Thompson [7]. Stylianides, Stylianides, and Philippou [14] conducted a study comparing participants' responses to conditionals posed in verbal (here, equivalent to the intuitive) and symbolic contexts. They found that the group who understood the underlying logical structures (mathematics majors) performed equally with the group who did not understand the logical structures (education majors) on the verbal context. However, the mathematics majors performed significantly better on the symbolic context. Hence, a next step in transfer from intuitive and abstract contexts to more applied contexts may be a symbolic context in which the intuitive nature of the context is removed, but with symbolic expressions as the content.

Counterintuitive context conditionals are defined as conditionals where the application of the reasoning is counter to the learner's prior knowledge [15-16]. These conditionals constitute a further step in the application of reasoning about conditionals. When this conflict exists, a learner must engage in what is known as "belief inhibition" to overcome the conflict and successfully follow the underlying logic to the conclusion of the conditional statement. Researchers De Neys and Franssens [15] conducted experiments which indicated that participants' reasoning about conditionals was negatively impacted when cued beliefs conflicted with the associated logic.

Comparison of Reasoning in Multiple Contexts

Recent research has focused on how individuals reason when conditionals are presented in one context as compared to another. Thompson [7] examined participants' reasoning about conditionals in three separate contexts: abstract (represented by the Wason Selection Task), intuitive (represented by social contract questions like those found in Cosmides & Tooby's study [6]), and factual (where statements were listed as facts, but no intuition was associated). Similar to Cosmides and Tooby's findings, participants most correctly answered the questions related to the intuitive context conditionals, but again generally failed to correctly respond to the contrapositive piece of the Wason Selection Task. This further supports the idea that individuals possess a natural understanding of conditionals in intuitive contexts and fail to transfer this understanding to conditionals in abstract contexts.

Christoforides et al. [8] compared intuitive, symbolic, and abstract contexts, and found that intuitive contexts made correct reasoning about the converse and inverse (generally difficult logical structures) easier for young children than when the contexts were abstract or symbolic. Several researchers have also examined conditionals as they are posed in the counterintuitive

context as compared with the intuitive context. For example, Vamvakoussi et al. [9] found that participants performed worse on counterintuitive conditionals than on conditionals that did not pose cognitive conflict (i.e., intuitive conditionals).

These studies indicate that reasoning, which is sufficient in an intuitive or abstract context, may not be sufficient when a learner encounters conditionals presented in the symbolic or counterintuitive contexts. Therefore, correct reasoning about more complex logical statements may be easier to achieve given an intuitive context. However, this reasoning may not transfer to other contexts due to interference from an individual's prior knowledge about the content of the logical statement or the context in which it is conveyed.

Transitioning from Context to Context

Transitioning learners from successfully reasoning in one context to another context requires an examination of what kind of reasoning is occurring in each context and how that reasoning may differ from reasoning in another context. Empirical understanding, that is understanding based on experience and observation of an idea, is not uncommon to the novice learner. Researchers have found that when novices attempt to prove a statement in a symbolic context or when they examine proof attempts (sample work that attempts to prove a conjecture, but does not constitute a complete or rigorous proof) they are very likely to accept empirical justifications as rigorous methods of proof [3, 17-20]. Experts, on the other hand, expect rigorous deductive arguments for any proofs or conjectures made within their subject area, but were willing to accept empirical justifications when presented with proofs that were outside of their domain [21, 22], reverting to the more simplistic reasoning used in intuitive contexts. These studies illustrate the idea that learners will revert to using an intuitive or empirical justification in

unfamiliar contexts.

Research also indicates that without the aid of an outside catalyst, learners will retain their empirical justifications as sufficient for a method of reasoning or proving in any context [3, 10]. Stylianides and Stylianides [3] facilitated this transition from less sophisticated empirical justification to more formal reasoning by presenting undergraduate mathematics education majors with a “monstrous counterexample.” The “monstrous counterexample” was the idea of a seemingly true mathematical formula whose counterexample was quite inconceivably out of the range of what an empirical justification would be able to prove or disprove, which created a sincere cognitive conflict for the learner to overcome. For example, researchers gave students the conjecture: “The expression $1 + 141n^2$ for n a natural number never equals a square number.” Students tested a few cases using their method of empirical justification and concluded that the conjecture was indeed true. Then researchers revealed that, while the conjecture was true for the natural numbers 1 through 30,693,385,322,765,657,197,397,207, the next natural number yielded a square number when inputted to the expression. This truly monstrous counterexample surprised the students and led them to conjecture that even their most intense forms of empirical justifications would not be feasibly sufficient and that they must begin looking for other more formal methods of reasoning.

Several researchers have proposed instructional progressions for reasoning about conditionals. For example, Stylianides and Stylianides [3] proposed a progression from less to more sophisticated reasoning and included the requirement of some sort of catalyst so that students would have sufficient “intellectual need” to transition from one level of sophistication to the next. Lee [23] presented his progression of proof construction within the symbolic context as

a continuum from irrelevant inferences to deductive proofs. In addition to suggesting that students must progress from informal to more formal types of reasoning, researchers also suggest that reasoning must be able to be completed in a variety of contexts, not just those in which the learners have some familiarity or form of intuition [24].

Each of these proposed progressions supports the idea that a novice learner will maintain empirical justifications regardless of context at the least sophisticated levels of a progression. To transition the learner from this type of reasoning, an outside catalyst of some form must be presented to move the learner to more formal reasoning. Overall, the research literature supports a larger progression where learners transition to more and more formal reasoning in each context and through each of the four contexts: intuitive, abstract, symbolic, and counterintuitive.

Methods

Based on the review of the literature, we hypothesized that the ordering of the contexts in which the conditionals are presented would influence students' reasoning about conditionals. An overarching research question and two sub-research questions guided the testing of this theory:

Overarching Research Question: How does the order of teaching four conditional contexts influence reasoning about conditionals?

Research Question 1: How does the order of teaching four conditional contexts using the Learning Logic App influence learners' performance on the Conditional Logic Assessment?

Research Question 2: What are learners' perceptions about how the order of teaching four conditional contexts influenced their performance on reasoning about conditionals?

Research Design

The research design for this study was a convergent parallel mixed methods design [25] that included the collection of quantitative and qualitative data. This design was chosen to provide a more complete understanding of participants' experiences with the different context

orderings through the complementarity of the quantitative and qualitative data [26].

The quantitative data included pre- and post-assessment scores on the Conditional Logic Assessment (CLA) and score logs from the Learning Logic App. The CLA is an assessment designed to examine a learner's reasoning about conditionals in multiple contexts. Qualitative data included participants' responses to a computer survey. This data source provided insights about participants' perceptions of how the context orderings influenced their performance on reasoning about conditionals.

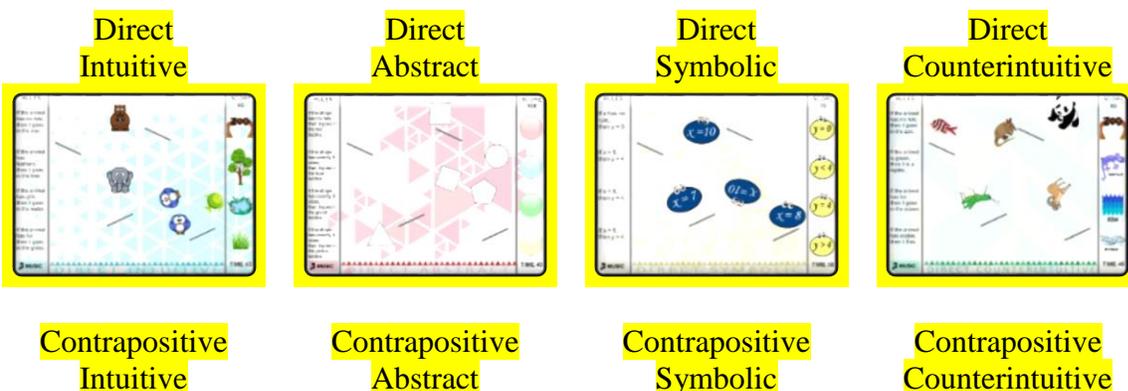
Participants

The participants in this study were 154 adults (over the age of 18 years) enrolled in an undergraduate program at a large public university in the intermountain west of the United States. Participants volunteered for the study through the university's SONA System, a system for recruiting undergraduate student participants. Participants were screened for prior experience with formally studying logic, such as in a university course. If participants had formally studied logic, they were excluded from the study. **There is a chance these participants may have informally encountered logic study, such as in a mathematics or computer science course. However, it is unlikely they would have examined conditionals as portrayed in all four of the contexts of this study.** The participants were randomly assigned to one of 23 context orderings for their interaction with the Learning Logic App.

Learning Logic App

The treatment used in this study was an educational app, called the Learning Logic App. The Learning Logic App employs a simply designed sorting game to expose players to conditionals in the four contexts: intuitive, abstract, symbolic, and counterintuitive. The rules of

the game require players to reason about conditionals in each of these contexts with two different logical structures: direct reasoning and contrapositive reasoning. Figure 1 depicts each of the four contexts and two types of reasoning in the app. The intuitive context involved reasoning about rules for familiar animals and their habitats. The abstract context involved reasoning about rules for shapes and colors. The symbolic context involved reasoning about rules with simple inequalities and numbers. The counterintuitive context involved reasoning about animals and following rules counterintuitive to those expected for each animal. The scoring for the app involves a single increment or decrement for each correct or incorrect item sorted. Additionally, a score decrement may also occur when the time available to sort an item elapses. Each change in score is recorded by a C# script into a score log described further below. The Learning Logic App was designed based on the research about conditional reasoning and educational app design, such as incorporating “low floors” [27] and simultaneous linking affordances [28, 29]. There are 24 possible permutations of the four contexts in the Learning Logic App, but 23 possible permutations were used. The Symbolic-Abstract-Counterintuitive-Intuitive (SACI) ordering was excluded as a condition because it follows the same context ordering as the Conditional Logic Assessment and may have caused a carryover effect [30].



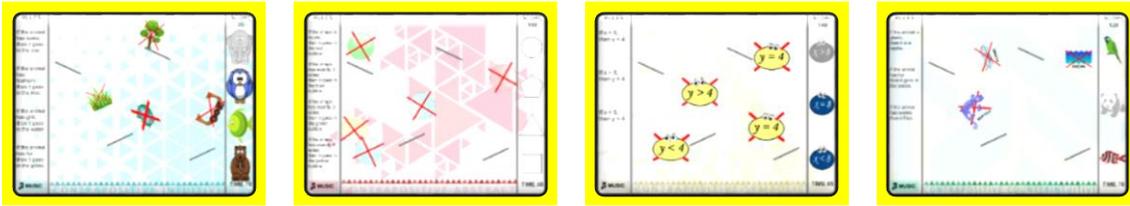


Figure 1. Screenshots of the four contexts and two reasoning types of the Learning Logic App.

Data Sources

There were three sources of data used in this research: Conditional Logic Assessment, score logs, and participant responses. The sections below include descriptions of each data source.

Conditional Logic Assessment. The Conditional Logic Assessment (CLA) is an assessment designed to assess an individual's reasoning about conditionals in different contexts and with different logical structures. The CLA includes 32 questions that are a composite of questions from the research literature on conditionals. There are three sections on the CLA that present conditionals in four different contexts: intuitive, abstract, symbolic, and counterintuitive. The CLA was used to measure change in reasoning about conditionals for each of these four contexts.

The CLA was developed by the primary author of this paper. Intuitive questions included six questions from Cosmides and Tooby's [6] research on conditionals in intuitive contexts. Participants' prior knowledge about the given scenario are likely to have a positive influence on their ability to correctly answer these questions. For example, in a scenario regarding alcohol consumption laws, "A person is sitting at the bar drinking water. Do you need to check this person's age to see if they are following the law? Yes, check this person's age. OR No, do not check this person's age." The correct solution here is "No, do not check this person's age."

Abstract questions were drawn from Wason [12] and Watts and Esty [31]. Four questions were derived from the Wason Selection Task [12]. Four abstract questions were drawn from Watts and Esty [31] with conditional statement questions using nonsense words but following the conditional “If H, then C” structure. In each of the abstract context questions, participants’ prior knowledge did not assist them in correctly answering the questions.

Symbolic questions were drawn from Watts and Esty [31]. For example, “Suppose this statement is true: ‘If $k > 4$ then, $j > 12$ ’. If ‘ $j = 6$ ’ is also true, what can be deduced?” The correct answer here is “ k is less than or equal to 4.” The ability to reason about these questions should not have been improved by an individuals’ prior knowledge but may be confounded by it.

Eight counterintuitive context questions were derived from De Neys and Van Gelder [16] and De Neys and Franssens [15]. These conditionals included intuitive contexts and counterintuitive contexts. In the counterintuitive context, participants’ prior knowledge may inhibit their ability to correctly answer these questions. For example, “Suppose this statement is true: ‘All vehicles have wheels.’ If ‘A boat is a vehicle’ is also true, can the following statement be deduced? ‘A boat has wheels.’ Yes, the statement can be deduced. OR No, the statement cannot be deduced.” The correct answer here is “Yes, the statement can be deduced.”

Score logs. The second source of data was score logs generated by a C# script in the Learning Logic App. The score logs recorded every change in a player’s score as well as the time the change occurred and the cause for the change in score. These data revealed the small shifts in understanding exhibited by the participants [32] and whether participants were actively playing with the Learning Logic App throughout the entire interaction period.

Participant survey. Participants completed a computer-based survey developed for this

study to provide feedback about their experience and on the app. In the survey, participants were asked their perceptions about four aspects of the different context orderings: their opinion of the order in which they experienced the contexts, a rating (on a continuum) of how the order helped or hindered their learning, a ranking of the contexts from easiest to hardest, and how their experience with the app influenced their performance on the post-assessment.

Procedures

Participants completed the Conditional Logic Assessment as a pre-assessment. They were then randomly assigned to one of 23 context orderings to interact with the Learning Logic App for approximately 15 minutes. Next, participants completed the Conditional Logic Assessment as a post-assessment. Finally, participants responded to a computer survey about their perceptions of the context ordering and their awareness of features in the app.

Data Analysis

There were three sources of data for the analysis: pre and post CLA, score logs and participant surveys. The quantitative data analysis of the CLA included descriptive statistics, visual and graphical analyses, and a multiple regression. Before running the regression, data were tested for linear relationships, multicollinearity, and outliers to ensure the assumptions of the regression were met. The independent variables were the pre-assessment score and the 23 context orderings. The context orderings were treated as a categorical variable. The dependent variable was the post-assessment score. The multiple regression allowed for prediction of the post-assessment score by both the pre-assessment score and the context orderings.

Researchers exported the score log data into a Microsoft Excel spreadsheet to be cleaned for analysis. In the spreadsheet, the columns denoted time and score. Each participant had two

columns: one for the time of the interaction that caused the change in score and one for the score. We plotted the data in line graphs for a visual analysis and comparison with other participants. This enabled us to examine the score log data in its entirety [33]. The score logs were a means of maintaining internal validity. Because participants completed all parts of the study anonymously and online, the score logs provided a way to “watch” participants’ interactions with the Learning Logic App. Therefore, anomalies, such as the participant who encountered technical difficulties and was unable to interact with levels 5-8, were recorded in the score logs and identifiable during the analysis. This mitigated the potential effects of active app engagement as a confounding variable.

The participant surveys were recorded in Qualtrics. We exported these data to NVivo for analysis. First, data from the surveys were open coded [34]. This was followed by axial coding to identify salient themes in the responses related to the conditional ordering in which the participants were assigned [35]. For the question involving ratings, participants were able to select their rating on a continuum for each question. Consequently, the researchers examined the ratings for each context ordering and then averaged them to provide a descriptive statistic. Finally, researchers synthesized and summarized the surveys. These data aided in triangulation and validation of the quantitative results [36].

Results

The results presented below include a quantitative analysis of the CLA, a visual analysis of the score logs, and a qualitative analysis of participants’ perceptions based on their survey responses. A total of 157 participants completed the pre and post Conditional Logic Assessments, played the Learning Logic App, and responded to the computer survey. The score

logs revealed that three participants allowed time to elapse and did not interact with the Learning Logic App at all; they were excluded from the data analysis. Therefore, the final data analyses include a total of 154 participants.

Quantitative Results

The quantitative results include descriptive statistics of the CLA scores, a multiple regression of the CLA scores, and a visual analysis of the score logs to address Research Question #1.

Descriptive statistics from the CLA. The first analysis was a visual inspection of the data for observable trends. Figure 2 shows a summary dot plot of the average change from pre- to post-assessment for each of the 23 context orderings. The horizontal axis of this graph demarks the 23 context orderings. The context orderings are arranged from lowest mean change to highest mean change from pre- to post-assessment. The error bars on the graph represent one standard deviation.

For all participants, there was an average change of 0.92 (indicated by the dashed line in Figure 2). Notably, only three context orderings had an average change greater than two (SIAC: $M = 3.6$, ASCI: $M = 3.0$, CIAS: $M = 2.14$). Four context orderings had a negative average change (see red squares in Figure 2; CSAI: $M = -1.14$, ICSA: $M = -0.71$, ACSI: $M = -0.57$, SICA: $M = -0.33$). Examining the error bars of the plot, only four context orderings had a positive change when considering one standard deviation less than the mean (see green circles in Figure 2). All other context orderings (indicated by blue triangles and red squares) drop below zero when considering one standard deviation below the mean.

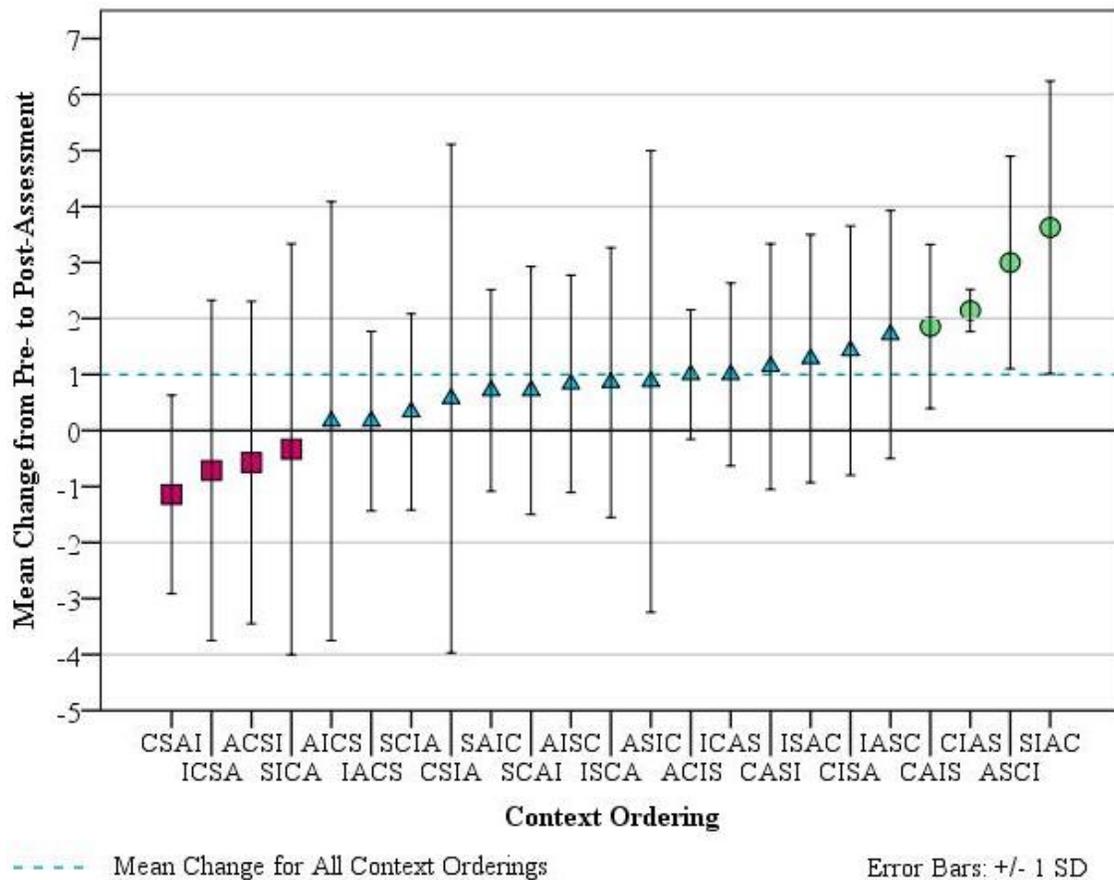


Figure 2. Summary dot plot of the change from pre- to post-assessment organized by mean change for all context orderings with one standard deviation error bars. The blue dotted line indicates the mean change for all context orderings. The red squares indicate context orderings which showed a decrease on average, the blue triangles indicate context orderings which showed a decrease on average, with one standard deviation dipping below zero; and the green circles indicate context orderings which showed an increase on average with one standard deviation still above 0.

The mean and standard deviation for the pre-assessment scores, post-assessment scores, and change between the pre-assessment and post-assessment for each context ordering are provided in Table 1. The contexts are ordered from the lowest to the highest average change. As can be seen in Table 1, the average pre-assessment score ranged between 17.17 and 22.86, the average post-assessment score ranged between 17.33 and 24.14, and the average change score ranged between -1.14 and 3.63.

Table 1

Descriptive Statistics from the Conditional Logic Assessment for the Pre-Assessment, Post-Assessment, and Change Scores

Context ordering	Prescore		Postscore		Change Score	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
CSAI	22.42	4.50	21.29	3.88	-1.14	1.64
ICSA	20.71	2.61	20.00	3.89	-0.71	2.81
ACSI	19.57	1.59	19.00	1.50	-0.57	1.09
SICA	20.67	1.89	20.33	2.42	-0.33	1.50
AICS	20.17	2.14	20.33	1.87	0.17	1.60
IACS	17.17	2.27	17.33	2.61	0.17	0.65
SCIA	19.50	1.20	19.83	1.38	0.33	0.72
CSIA	19.71	2.17	20.29	1.90	0.57	1.72
SAIC	20.14	2.05	20.86	1.82	0.71	0.68
SCAI	20.00	1.36	20.71	0.81	0.71	0.84
AISC	21.67	2.06	22.50	2.26	0.83	0.79
ISCA	18.14	1.93	19.00	1.60	0.86	0.91
ASIC	19.13	2.07	20.00	1.76	0.88	1.46
ACIS	16.86	2.24	17.86	2.04	1.00	0.44
ICAS	19.00	3.70	20.00	3.34	1.00	1.51
CASI	19.14	1.68	20.19	1.64	1.14	0.83
ISAC	22.86	2.04	24.14	2.13	1.29	0.84
CISA	17.86	2.56	19.29	2.27	1.43	0.84
IASC	17.43	1.69	19.14	1.28	1.71	0.84
CAIS	20.43	1.73	22.29	1.87	1.86	0.55
CIAS	17.57	1.13	19.71	1.21	2.14	0.14
ASCI	18.33	1.36	21.33	1.17	3.00	0.78
SIAC	19.38	1.92	23.00	1.49	3.63	0.93
All orderings	19.48	4.77	20.47	4.53	0.99	2.52

N = 154.

Multiple regression of the CLA and context ordering. Next, we used a multiple regression to model participants' CLA post-assessment scores based on their pre-assessment scores and randomly-assigned context orderings (dummy-coded). In the dummy-coding of the context orderings, the CSAI ordering was designated as the reference category due to it being

hypothesized as the least beneficial context ordering. This context ordering and pre-assessment score combine to account for 73.7% of the variance of the post-assessment score, $R^2_{\text{adj}} = .737$, $F(23, 130) = 19.68$, $p < .001$. To ensure that no severe violations to the multiple regression assumptions occurred, researchers examined the residual diagnostic plots. There was no evidence of violation of the assumptions of normality of variance, homogeneity of variance, or outliers.

Two context orderings and the pre-assessment score displayed significant effects on the post-assessment score (pre-assessment: $b = 0.81$, $\beta = 0.86$, $p < .001$; SIAC ordering: $b = 3.81$, $\beta = 0.19$, $p = .003$; ASCI ordering: $b = 2.99$, $\beta = 0.13$, $p = .029$). No other context orderings resulted in statistically significant effects. Regression coefficients and p values for all ordering conditions from the model are provided in a table in Appendix A. This indicates that each additional question answered correctly on the pre-assessment (an increase of one unit in the pre-assessment score) is associated with an increase of 0.81 units on the post-assessment. Because prior knowledge should not have been hindered by the Learning Logic App interaction, researchers anticipated this positive influence of the pre-assessment on the post-assessment.

Holding the pre-assessment scores constant, the SIAC ordering was associated with a 3.81-point increase in a participant's post-assessment score compared to the CSAI ordering. That is, it was expected that participants in the SIAC condition would correctly answer three to four more questions on the post-assessment than those participants in the CSAI ordering. In the ASCI ordering, the model predicts participants will correctly answer three more questions on the post-assessment ($b = 2.99$) than participants in the CSAI ordering. When examining both of results in terms of the standardized coefficients (SIAC: $\beta = 0.19$, ASCI: $\beta = 0.13$), the effects are considered medium and small, respectively [37]. The results of the multiple regression indicate

that participants' prior knowledge (as measured by the pre-assessment) strongly influenced the post-assessment scores. In addition, the SIAC and ASCI orderings were the most positively influential on participants' post-assessment scores. It should be noted that the number of participants in each context ordering may have limited the power of the regression analysis.

Visual analysis of the score logs. The score logs recorded changes in each participant's score while they played each level of the Learning Logic App. The level order of the Learning Logic App depended upon the participants' random assignment to one of the 23 context orderings. Each score log plots the participants' score against time. Every level began with a score of 100 points and lasted for 100 seconds. Figure 3 presents four representative score logs which illustrate the four trends found across all the context orderings.

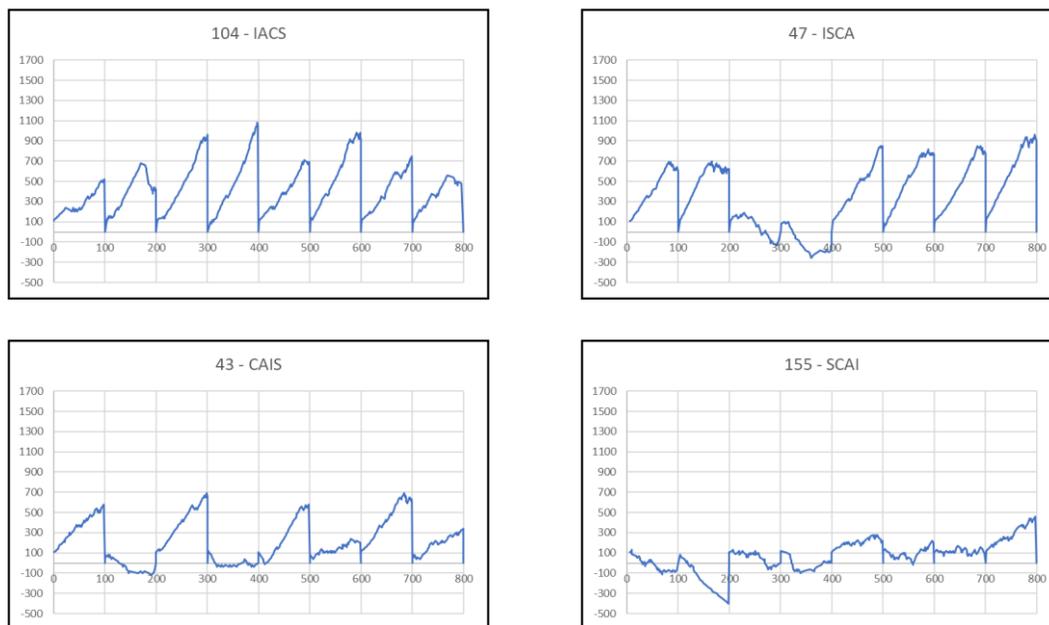


Figure 3. Representative score logs of four participants. Each 100 second section represents a different level of the Learning Logic App. The context ordering is given at the top of each graph.

The top left score log of Figure 3 depicts a participant who experienced very little

struggle while interacting with the app. A large portion of participants (39%) had similar experiences while interacting with the app. These participants exhibited their understanding by ending each level with high scores, regardless of context or task. The top right score log of Figure 3 depicts a participant who had considerable difficulty with the symbolic levels (levels 3 and 4) as compared with levels in the other contexts. This is the first time the participant would have encountered the most challenging symbolic context. This pattern appeared in 26% of participants. Interestingly, this pattern is unique to the symbolic context. The bottom left score log of Figure 3 depicts a participant who performed similarly across all contexts, but struggled with the contrapositive reasoning (levels 2, 4, 6, and 8). Other participants who exhibited similar patterns (12%) struggled with the contrapositive reasoning in one, two, or three of the levels while still performing well on the direct reasoning. Finally, the bottom right score log of Figure 3 depicts a participant who experienced a large amount of struggle throughout all eight levels with a little bit of improvement towards the final two levels (intuitive context). Few participants exhibited this level of struggle for all levels (7%). These participants struggled on all levels, regardless of context or task. The remaining 16% of participants exhibited patterns which included multiple trends described above (e.g., difficulties on the symbolic context and the contrapositive reasoning) or exhibited no discernable pattern at all.

Qualitative Results

The qualitative results focus on participants' responses on the online survey that they completed after the post-assessment. The survey asked participants about their perceptions of the different conditional contexts.

Perceptions of different context orderings and contexts. Participants reported their

perceptions about four aspects of the different context orderings: opinion of the order they experienced the contexts, a rating of how the order helped or hindered their learning, a ranking of the contexts from easiest to hardest, and how their experience with the app influenced their performance on the post-assessment. Approximately two-thirds of the participants felt that the order of the levels they experienced was good (50%) or okay (15%), while 13% of participants felt the order should be changed. Interestingly, the only context ordering where all participants thought the order was “good” and most (75%) felt the order was “progressive” was the IACS ordering (intuitive-abstract-counterintuitive-symbolic). This ordering was very similar to the proposed ordering IASC. One participant in the IACS ordering said, “I thought they were presented in an order that was effective and that built upon each other.” Another IACS participant said, “It got tougher and tougher each level, which was perfect.” In contrast, most participants (70%) in the SCIA and the CSIA context orderings felt that the order of contexts “didn't flow as well as it could have” or that “they were in a weird level [order] that got a little easier every time.” These were context orderings where the two most challenging contexts, symbolic and counterintuitive, came first and second. A participant in the ASIC ordering (which is somewhat like the proposed ordering) suggested the order be changed:

I think intuitive would have been nice to have first because it's something, obviously, we should be familiar with while I have a hard time recalling rules for colors and shapes at such a high pace and it left me a little frustrated.

In addition to their written responses, participants rated how the context ordering impacted their learning (0 = very hindering, 5 = not helpful or hindering, 10 = very helpful). Across all context orderings, the participants rated the ordering of the levels as 6.19, which would be considered a little helpful to their learning on the given scale. Once these results were

considered by the context orderings, the results were more distinctive. The IACS ordering participant group rated the ordering of the level the most helpful out of all the context orderings (7.67). This aligns with the IACS participants written responses about the order of the four contexts which indicated that participants felt the context ordering was good and progressive. The average ratings from the CAIS, AICS, CIAS, and ICSA participant groups (a mix of context orderings) were not far behind (7.57, 7.4, 7.14, and 7, respectively). Participants in the ICAS, CSAI, and CASI context orderings rated the level ordering as a 5 (neutral to their learning) on average. The only context ordering with an average rating indicating that the participants felt their learning was hindered by the ordering was CSIA (4.86).

Participants also ranked the difficulty of the contexts from easiest to most difficult (1 = easiest to 4 = hardest). Overall, participants consistently rated the symbolic context as the most difficult (average rating of 3.48) and the intuitive context as the easiest (average rating of 1.86). The participants rated the abstract and counterintuitive contexts nearly equally with the abstract context slightly more difficult (average rating of 2.42) followed by the counterintuitive context (average rating of 2.24). When examining these results by context ordering, participants in all context orderings but one (CIAS) indicated that the symbolic context was the most difficult. Participants in two thirds of the context orderings also indicated the intuitive context was easiest. There were mixed results as to the comparative difficulty of the abstract and counterintuitive contexts.

When asked if they felt the Learning Logic App influenced their responses on the post-assessment, 51% of participants responded “yes,” 32% responded “no,” and 13% said “it influenced some questions.” Participants in 14 of the 23 context orderings felt that the app

influenced their responses on the post-assessment. In the ASCI ordering, participants had the most positive responses, with all participants indicating that the app influenced their responses on the post-assessment. A participant in the IACS ordering (in which 83% of participants said yes) responded, “Yes! I noticed I understood ALL the questions in full and changed a couple answers.” In contrast, participants in five of the context orderings (AISC, CIAS, ICAS, ICSA, and SAIC) had mixed responses about the influence of the app on their post-assessment, with responses such as “a little bit,” “maybe,” or “not really.” In the remaining three context orderings (CISA, CSIA, and ISAC), participants expressed that the app did not influence their post-assessment performance. One participant indicated that she “tried to implement it, but I think my answers were the same as before.”

Across these four aspects pertaining to participants’ perceptions of the context orderings and contexts, there were three major themes. First, the participants positively perceived the IACS ordering. Second, the participants consistently perceived the symbolic context as the most challenging. Third, participants consistently perceived the CSIA ordering as hindering to their learning or not influencing their responses on the post-assessment. Interestingly, the IACS ordering is very similar to the proposed context ordering, IASC. In contrast, the poorly perceived CSIA ordering is quite dissimilar.

Perceptions of the four context orderings with the highest average change. The four context orderings with the highest average change from pre- to post-assessment are shown as green circles in Figure 2. Participants in these context orderings had a mix of perceptions. These four context orderings, in order of highest to lowest average, were SIAC, ASCI, CAIS, and CIAS. Participants who experienced the SIAC ordering had positive perceptions of the context

ordering and the app's influence on their understanding. One SIAC participant said, "It made me think in a different way, it gave me a challenge." Another SIAC participant stated, "It challenges me to think about what the rule was in that first round [direct reasoning task], but then in the second round [contrapositive reasoning task] to recognize which rule it would follow without the x through it. I just thought it was a fun challenge." The SIAC ordering was the only context ordering of these four context orderings where the participants had consistent perceptions, and those perceptions were positive.

The participants who experienced the ASCI ordering had a range of perceptions with some participants expressing frustration to others who felt like they learned quite a bit. For example, one ASCI participant said, "I didn't like how it messed with logic" and "It was very frustrating." In contrast, another ASCI participant said, "I was impressed with what I learned just in a short amount of time." The perceptions of the other ASCI participants were similarly dispersed and did not reveal any consistent theme that was either positive or negative. The other two context orderings, CAIS and CIAS were not significantly different from the CSAI ordering, according to the multiple regression. Similar to the ASCI ordering, participants who experienced the CAIS ordering and the CIAS ordering had a wide range of perceptions of the context orderings and their influence on the participants' understanding. For the participants who experienced the CAIS ordering, some participants felt the ordering was beneficial while others felt it was just confusing. Similar to the ASCI and CIAS orderings, the participants assigned to the CIAS context ordering had inconsistent perceptions across the group. Although these four context orderings produced the highest performance results, the SIAC ordering was the only context ordering which had a consistent theme among the participants' perceptions.

Mixed Methods Results

This section discusses the meta-inferences derived from the quantitative (participants' performance) and qualitative (participants' perceptions) results to address the Overarching Research Question on how the order of teaching four conditional contexts influences reasoning about conditionals. First, we present the convergent results (i.e., results that appeared in both the quantitative and qualitative results), followed by the divergent results (i.e., results that appeared in only the quantitative or the qualitative results).

Convergent results. There are three meta-inferences derived from the quantitative and qualitative results. The first meta-inference relates to the positive impact of the SIAC ordering on learners' performance on and perceptions of reasoning about conditionals. Participants in the SIAC ordering had significant improvements on the CLA. Additionally, SIAC ordering participants highly rated the Learning Logic App overall and their enjoyment of the app. These indicators, from both the quantitative and qualitative results, suggest a positive benefit of the SIAC ordering for teaching reasoning about conditionals.

The second meta-inference concerns the negative influence of the CSAI ordering on participants' performance and perceptions. Overall, CSAI participants performed the worst, with an average negative change from pre- to post-assessment. CSAI participants' perceptions of the Learning Logic App indicated that the ordering was the worst at engaging the participants as compared to all other context orderings. These parallel results point to a potential inadequacy in the CSAI ordering for teaching reasoning about conditionals.

The final meta-inference regards the challenges participants encountered with the symbolic context. The score logs of the Learning Logic App clearly show participants struggling

with the symbolic context levels when they did not struggle with the levels for the other three contexts. Participants' awareness of their difficulties with the symbolic context were exhibited through their consistent ratings of the symbolic level as the most difficult. Additionally, participants' dislike of the symbolic context appeared as a theme in response to the survey questions about the "least favorite parts of the Learning Logic App."

Divergent results. There were two notable results that diverged in the quantitative and qualitative results. The first divergence reflects differing outcomes in the qualitative and quantitative results for the IACS ordering. For this context ordering, a disparity emerged between participants' strong praises for the IACS ordering and the mediocre to poor performance of the participants on the post-assessment. Participants in the IACS ordering had the most positive perceptions of their context ordering, the app overall, their enjoyment of the app, and how well the app kept their attention when compared with all other groups. However, when examining the average change in performance for IACS participants, the results indicated nearly a zero change. Thus, while participants liked the IACS ordering, it did not appear to benefit their reasoning about conditionals.

The second divergent result pertains to the influence of the ASCI ordering on the participants' performance and perceptions. While participants assigned to the ASCI ordering had significant improvements from pre- to post-assessment, their perceptions were varied. Participants in this context ordering had the highest rating of the app influencing their performance on the pre-assessment. However, these same participants had a variety of responses (both positive and negative) on all other questions related to understanding, specifically the question pertaining to participants' perceptions of the ordering. Consequently, participants'

performance was not aligned with their self-reported perceptions.

The meta-inferences derived from merging the quantitative and qualitative results suggest that the SIAC ordering may be the most beneficial context ordering for learners wanting to improve their reasoning about conditionals. The results show that learners may need more assistance or practice with the symbolic context than the other three contexts when learning about conditionals. The results are especially important when considering the order in which the contexts should be taught, the amount of instructional time that should be dedicated to each context, and the potential transfer of conditional reasoning in all four contexts to other areas both within and outside of mathematics.

Discussion

The purpose of this study was to examine how the order of teaching four conditional contexts influenced reasoning about conditionals. Here we discuss the four of the conditional contexts, the influence of context order on learners' performance and perceptions, and recommendations for education and future research.

Four Conditional Contexts

Each of the four conditional contexts (intuitive, abstract, symbolic, and counterintuitive) involved the same underlying logical structures. However, learners' performance and perceptions differed greatly across the four contexts. The results indicate that the symbolic context was the most difficult for learners, who struggled with this context on the pre- and post-assessments. The score logs also revealed participants struggling with the two symbolic context levels, regardless of if they encountered this context at the beginning, middle, or end of their interaction with Learning Logic App. Participants repeatedly expressed that the symbolic context was the most

difficult of the four contexts. This aligns with previous research studies comparing learners' performance in these contexts [7-8,14,38].

However, based on the literature, researchers did not expect the symbolic context to be more difficult for learners than the counterintuitive context. There are several possible explanations for this result. First, as suggested by De Neys and Van Gelder [16], undergraduate participants in this study are at the stage of life when belief inhibition (the skill necessary for reasoning in the counterintuitive context) comes most naturally, meaning that they may have found the counterintuitive context more difficult than the symbolic due to their decreased difficulties with belief inhibition. Second, as suggested by Mutodi [39], who found that "learners stick to procedurally driven symbols at the expense of conceptual and contextual understanding" (p. xi), the participants in the present study may have had very influential prior notions about the symbols used in the app and in the assessments, which conflicted with how the symbols were used in the symbolic context questions. Consequently, participants would need to mentally maneuver through the intended intricacies of the symbolic context as well as employ their skills of belief inhibition (which are required by the counterintuitive context) within one context. Based on this result, we recommend that more instructional time be allocated for the symbolic context as learners will most likely struggle with conditionals posed in this context.

Influence of Context Order on Learners' Performance and Perceptions

The focus of this study was the influence of context orderings on learners' reasoning about conditionals. The sections below discuss the context orderings that were most influential for the participants, both positively and negatively.

Symbolic-intuitive-abstract-counterintuitive ordering (SIAC). The results indicated

that the symbolic-intuitive-abstract-counterintuitive (SIAC) ordering positively influenced learners' reasoning about conditionals with the SIAC ordering showing significant improvements from pre- to post-assessment. Participants in this context ordering highly rated their experience with the Learning Logic App, which indicated positive perceptions of the context ordering. While this context ordering does not align with the initially proposed IASC ordering, it is similar to the IASC ordering with the most difficult context (symbolic) moved to the beginning (see Figure 4).



Figure 4. Comparison of the hypothetical IASC ordering and the SIAC ordering.

Because of participants' difficulty with the symbolic context, it is possible that the symbolic context acted as a catalyst for participants prompting them to reevaluate their conceptions of the conditionals (see Figure 4). The idea of a catalyst to prompt conceptual re-evaluation stems from Stylianides and Stylianides' study [3] where researchers attempted to sway students from purely empirical justifications to deductive reasoning by using a catalyst, called a "monstrous counterexample." In the present study, when participants interacted with the symbolic context levels first, their struggle with the symbolic level may have had a catalyzing effect encouraging them to reevaluate their underlying conceptions about how to reason about conditionals. Based on this result, we recommend that, to effectively influence a learner's reasoning about conditionals, it may be beneficial to use some form of catalyst, such as the symbolic context, at the beginning of an instructional intervention about conditionals to prompt

the learner to re-analyze prior conceptualizations. While there were no significant performance results for context orderings beginning with the counterintuitive context, future research could examine what contexts could act as the most effective catalyst for learners.

Counterintuitive-symbolic-abstract-intuitive ordering (CSAI). In contrast to the results for the SIAC ordering, the counterintuitive-symbolic-abstract-intuitive (CSAI) ordering negatively influenced learners' reasoning about conditionals, with an average decrease from pre- to post-assessment. Participants in the CSAI ordering reported that they had a difficult time keeping their attention focused while playing the Learning Logic App. Several expressed annoyance or confusion after interacting with this context ordering. Based on the literature, we proposed that the context ordering of intuitive-abstract-symbolic-counterintuitive (IASC) would be the most beneficial to participants, and that the reverse (CSAI, see Figure 5) would be the least beneficial.



Figure 5. Comparison of the hypothetical IASC ordering and the CSAI ordering.

The decrease from pre- to post-assessment for the CSAI participants could have been caused by their interaction with the Learning Logic App influencing their reasoning about conditionals. Perhaps participants became frustrated or fatigued with the app and subsequently performed poorly on the post-assessment. It is also possible that beginning with the counterintuitive context, followed immediately by the most difficult symbolic context, did not allow the learners to apply their newly disrupted reasoning pattern to the more challenging

contexts (counterintuitive and symbolic). Participants' confusion could have prevented them from developing new conceptualizations of reasoning about the conditionals.

Because this context ordering is the exact reverse of the proposed progression of IASC, using the CSAI ordering conflicts with underlying theories about conditional ordering. As Clements and Sarama [40] noted, progressions are designed with "levels of increasing sophistication, complexity, abstraction, power, and generality" (p. 83). The underlying theory of developmental progressions stems from the idea that "a critical mass of ideas from each level must be constructed before thinking...becomes ascendant in the child's mental actions and behavior" [41, p. 668]. The CSAI ordering presents decreasing levels of sophistication, complexity, abstraction, power, and generality and consequently does not allow for the accumulation of ideas from the less complex levels. We recommend that the CSAI ordering not be utilized in teaching reasoning about conditionals because it does not follow a developmental progression that increases in difficulty.

Intuitive-abstract-counterintuitive-symbolic ordering (IACS). In contrast to both the SIAC and CSAI ordering results, the performance and perceptions of participants in the intuitive-abstract-counterintuitive-symbolic (IACS) ordering were notably divergent. While IACS participants reported the most enjoyment and attentiveness of all the context orderings, these participants showed almost no gains from pre- to post-assessment. Participants described the context ordering as "progressive," feeling that it went from easy to hard. These results mean that, while participants in this context had very positive perceptions of the IACS ordering, their performance did not exhibit positive gains. Interestingly this ordering closely resembles the proposed IASC ordering but with the perceived most difficult context (symbolic) placed last (see

Figure 6).



Figure 6. The hypothetical IASC ordering and the IACSCS ordering.

Prior research has shown that many learners have misconceptions when it comes to reasoning about conditionals (e.g., [3-4, 10, 13]). One possible explanation for that lack of gains between the pre- and post-assessment for the IACSCS ordering is that it may not assist learners in correcting their misconceptions (see Figure 6) because it does not provide enough cognitive disruption to correct prior incorrect reasoning. Consequently, learners may rely on prior conceptualizations with support from the easier contexts at the beginning of the ordering. Then when reaching the difficult contexts at the end of the ordering, participants have had no opportunity to practice and receive confirmation for changes to their conceptions.

Chen, Schneps, and Sonnert [42] found a similar result when studying a simplified and then a realistic model of the solar system. Researchers found that students exposed to the realistic model first, and then the simplified model, made gains from each model. However, students exposed to the simplified model first, and then the realistic model, did not make gains from the second (realistic) model. Researchers posited that, when introduced to the simplified model first, the simplified model “which requires less cognitive load, anchors students’ understanding, and they appear reluctant to change their conceptualization when exposed to a model that requires a higher cognitive load” [42, p. 815]. For the IACSCS ordering participants, the initial low cognitive load of the intuitive context may have re-anchored participants’ understanding in their prior

knowledge, making them reluctant to change their conceptualization when encountering the higher cognitive load of the counterintuitive and symbolic contexts. While this context ordering is progressively complex and difficult, it fails to provide sufficient influence on the learner's performance in reasoning about conditionals.

Recommendations for Researchers and Educators

Based on participants' performance gains with the SIAC ordering and participants' positive perceptions of the IACS ordering, we recommend consideration of a new context ordering for researching and teaching reasoning about conditionals. This new context ordering is: symbolic-intuitive-abstract-counterintuitive-symbolic (SIACS, see Figure 7).



Figure 7. Proposed symbolic-intuitive-abstract-counterintuitive-symbolic (SIACS) context ordering for teaching reasoning about conditionals.

This context ordering incorporates the catalytic effects of the symbolic context being presented first and the progressive difficulty of the IACS ordering presented immediately thereafter. This proposed ordering capitalizes on the positive potential influences for learners' performance and perceptions of the conditionals. We also recommend careful allocation of instructional emphasis with extra time reserved for students to interact with the symbolic context. The persistent difficulties participants in this study encountered with the symbolic context indicate that the equal emphasis given to all contexts in the design of the Learning Logic App did not allow sufficient time for the participants to reform their conceptualizations or did not

effectively change their symbolic conceptualizations at all. Consequently, we recommend that the app design be modified to include features, such as increased exposure time, that provide a focused constraint affordance regarding the reasoning involved within the symbolic context.

Conclusion

This study examined how the order of four conditional contexts influenced learners' reasoning about conditionals. The mixed methods examination of learners' reasoning through the perspectives of performance and perceptions allowed researchers to measure change in performance and to make inferences about learners' shifts in conceptualization of reasoning about conditionals. The results showed **that the SIAC ordering** had the most positive influence on learners' reasoning and the most positive learner perceptions. This finding, as related to the literature on conditionals, led to the suggestion of a new context ordering: symbolic-intuitive-abstract-counterintuitive-symbolic (SIACS). We hypothesize that this new ordering may be more beneficial to learners because of the catalytic effects of the difficult symbolic context presented at the beginning of the ordering and revisited later at the end of the ordering.

These findings advance the research literature on reasoning about conditionals in different contexts and in transitioning learners' from correctly reasoning in one context to the next. The results provide educators with a more efficient instructional program to improve learners' reasoning about conditionals and aid learners to become citizens who can interpret and capitalize on the conditionals they encounter in everyday life.

References

1. Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 805-842). Charlotte, NC: Information Age Publishing.
2. Zandieh, M., Roh, K. H., & Knapp, J. (2014). Conceptual blending: Student reasoning when proving “conditional implies conditional” statements. *The Journal of Mathematical Behavior*, 33, 209-229.
3. Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, 40(3), 314-352.
4. O'Brien, T. C. (1974). Logical thinking in college students. *Educational Studies in Mathematics*, 5(1), 71-79.
5. Kılıç, D., & Sağlam, N. (2014). Students' understanding of genetics concepts: The effect of reasoning ability and learning approaches. *Journal of Biological Education*, 48(2), 63-70.
6. Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163-228). New York, NY: Oxford University Press.
7. Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, 76(3), 209-268.
8. Christoforides, M., Spanoudis, G., & Demetriou, A. (2016). Coping with logical fallacies: A developmental training program for learning to reason. *Child Development*, 87(6), 1856-1876. <https://doi.org/10.1111/cdev.12557>

9. Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2013). Brief report. Educated adults are still affected by intuitions about the effect of arithmetical operations: Evidence from a reaction-time study. *Educational Studies in Mathematics*, 82(2), 323-330.
<https://doi.org/10.1007/s10649-012-9432-8>
10. Artman, L., Cahan, S., & Avni-Babad, D. (2006). Age, schooling and conditional reasoning. *Cognitive Development*, 21(2), 131-145.
11. Herman, G. L., Loui, M. C., Kaczmarczyk, L., & Zilles, C. (2012). Describing the what and why of students' difficulties in Boolean logic. *ACM Transactions on Computing Education*, 12(1), 1-28.
12. Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, 20(3), 273-281.
13. Wagner-Egger, P. (2007). Conditional reasoning and the Wason selection task: Biconditional interpretation instead of reasoning bias. *Thinking & Reasoning*, 13(4), 484-505.
14. Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics*, 55(1), 133-162.
15. De Neys, W., & Franssens, S. (2009). Belief inhibition during thinking: Not always winning but at least taking part. *Cognition*, 113(1), 45-61.
16. De Neys, W., & Van Gelder, E. (2009). Logic and belief across the lifespan: The rise and fall of belief inhibition during syllogistic reasoning. *Developmental Science*, 12(1), 123-130.
17. Demiray, E., & Bostan, M. I. (2015). An investigation of pre-service middle school mathematics teachers' ability to conduct valid proofs, methods used, and reasons for

- invalid arguments. *International Journal of Science and Mathematics Education*, 15(1), 1-22.
18. İmamoğlu, Y., & Toğrol, A. Y. (2015). Proof construction and evaluation practices of prospective mathematics educators. *European Journal of Science and Mathematics Education*, 3(2), 130-144.
19. Lachmy, R., & Koichu, B. (2014). The interplay of empirical and deductive reasoning in proving “if” and “only if” statements in a dynamic geometry environment. *The Journal of Mathematical Behavior*, 36, 150-165.
20. Stylianides, A. J. (2007b). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 289-321.
21. Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, 431-459.
22. Weber, K. (2013). On the sophistication of naïve empirical reasoning: factors influencing mathematicians’ persuasion ratings of empirical arguments. *Research in Mathematics Education*, 15(2), 100-114.
23. Lee, K. (2016). Students’ proof schemes for mathematical proving and disproving of propositions. *The Journal of Mathematical Behavior*, 41, 26-44.
24. Babai, R., Levyadun, T., Stavy, R., & Tirosh, D. (2006). Intuitive rules in science and mathematics: A reaction time study. *International Journal of Mathematical Education in Science and Technology*, 37(8), 913-924.
25. Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

26. Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social & behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
27. Burke, Q., & Kafai, Y. B. (2014). Decade of game making for learning: From tools to communities. *Handbook of Digital Games*, 689-709.
28. Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments*, 4(3), 35-50.
29. Moyer-Packenham, P. S., & Westenskow, A. (2016). Revisiting the effects and affordances of virtual manipulatives for mathematics learning. In K. Terry & A. Cheney (Eds.), *Utilizing virtual and personal learning environments for optimal learning* (pp. 186-215). Hershey, PA: IGI Global.
30. Cohen, B. H. (2013). *Explaining psychological statistics* (4th ed.). Hoboken, NJ: Wiley.
31. Watts, C. M., & Esty, W. W. (2013). *Assessing conditional logic: Alternatives to Wason*. Unpublished manuscript, Department of Mathematical Sciences, Montana State University, Bozeman, MT.
32. Watts, C. M., Moyer-Packenham, P. S., Tucker, S. I., Bullock, E. P., Shumway, J. F., Westenskow, A., ... Jordan, K. (2016). An examination of children's learning progression shifts while using touch screen virtual manipulative mathematics apps. *Computers in Human Behavior*, 64, 814-828.
33. Dickinson, W. B. (2010). Visual displays for mixed methods findings. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 469-504). Thousand Oaks, CA: Sage.

34. Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13(1), 3-21.
35. Saldaña, J. (2015). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
36. Nastasi, B. K., Hitchcock, J. H., & Brown, L. M. (2010). An inclusive framework for conceptualizing mixed methods design typologies: Moving toward fully integrated synergistic research models. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 305-338). Thousand Oaks, CA: Sage.
37. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
38. Case, J. (2013). *Calculus students' understanding of logical implication and its relationship to their understanding of calculus theorems* (Doctoral dissertation). Orono, ME: University of Maine.
39. Mutodi, P. (2016). *Mathematical symbolisation: Challenges and instructional strategies for Limpopo Province secondary school learners* (Doctoral dissertation). University of South Africa, Pretoria, South Africa.
40. Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2), 81-89.
41. Sarama, J., Clements, D. H., Barrett, J., Van Dine, D. W., & McDonel, J. S. (2011). Evaluation of a learning trajectory for length in the early years. *ZDM Mathematics Education*, 43, 667-680.

42. Chen, C., Schneps, M. H., & Sonnert, G. (2016). Order matters: Sequencing scale-realistic versus simplified models to improve science learning. *Journal of Science Education and Technology*, 25(5), 806-823.

Appendix A

Multiple Regression Coefficients and Statistics

Variable	Standardized coefficient, b	Standardized coefficient, β	t	Significance, p
Assessment	0.813	0.858	19.526	.001***
Orderings				
	3.811	0.188	3.023	0.003**
	2.991	0.128	2.213	0.029*
	2.240	0.103	1.731	0.086
	2.124	0.098	1.646	0.102
	1.991	0.092	1.522	0.130
	1.536	0.071	1.173	0.243
	1.448	0.062	1.081	0.282
	1.330	0.061	1.018	0.310
	1.285	0.059	0.989	0.324
	1.044	0.048	0.806	0.422
	1.017	0.047	0.785	0.434
	1.014	0.050	0.804	0.423
	0.907	0.039	0.674	0.501
	0.821	0.038	0.633	0.528
	0.813	0.038	0.623	0.535
	0.761	0.033	0.567	0.572
	0.715	0.033	0.544	0.587
	0.543	0.023	0.403	0.687
	0.501	0.021	0.373	0.710
	0.094	0.004	0.070	0.944
	-0.061	-0.003	-0.045	0.964
	-0.349	-0.016	-0.269	0.789

* $p < .05$.
 ** $p < .01$.
 *** $p < .001$.