Pull the Andon rope on working memory capacity interventions until we know more

Sandra Gillam, PhD., CCC-SLP
Sarai Holbrook, MS., CCC-SLP
Jamie Mecham, MS., CCC-SLP

Daylene Weller
Utah State University

Correspondence about this paper should be sent to: Sandra L. Gillam, PhD; Language and Literacy Clinic; 2610 Old Main Hill; Logan, UT 84322 or to sandi.gillam@usu.edu.
Abstract

Purpose: The purpose of this article is to discuss the current state of interventions for improving working memory (WM) capacity language, and academic skills and to provide suggestions for speech language pathologists working with students who have WM capacity limitations.

Method: Meta-analyses, systematic reviews, randomized controlled trials and non-randomized comparison studies investigating the role of WM interventions for improving WM capacity language, and academic skills are reviewed. Strategies for improving WM are discussed.

Results: The use of interventions designed to improve WM capacity and other cognitive skills is currently not supported by the research. Direct working memory interventions should be considered to be experimental at this time. Such interventions require further investigation before they are used regularly for children with Developmental Language Disorders.

Discussion: Clinicians and practitioners should look to already established interventions for improving how students with DLD utilize organizational strategies and other well-researched methods for improving their cognitive and academic functioning in functional contexts.
In the 2013 book, *The Slow Fix*, author Carl Honore persuasively argued that our nation is addicted to quick fixes for solving complex problems. He pointed out that in many cases, this has resulted in the pursuit of short-term remedies over long-term solutions. The idea that slow fixes are best used to solve complex problems was described in the book, *Toyota Kata* by Mike Rother (2011). In his book he described how the Toyota automobile company implemented a program from 1950-1990 whereby the entire manufacturing line could be brought to a halt by one of the workers on the floor simply by pulling a rope. The rope was called an “Andon rope.” The idea was that any one of the skilled workers on the line had the power to identify that there was a problem that needed to be solved before the factory returned to operations. The result of this process was that for about 40 years Toyota was known for building cars that were unprecedented in quality, which led to success that, as far as we know, has never been replicated in the car business.

There are a number of examples of when we, as a society, have adopted quick fixes over pulling the Andon rope to carefully examine the evidence before moving forward. A recent example involves the use of Body Worn Cameras (BWCs) by police officers. The thought process behind asking officers to wear BWCs was based on a psychological phenomenon known as, “The Hawthorne Effect” which has shown that people alter their behavior when they are aware that they are being watched (Gillespie, 1991). This phenomenon contributed to the widespread use of BWCs by police all over the country (U.S. Department of Justice, 2015). However, a recent, large-scale randomized trial of over 2,200 officers in Washington D.C., revealed no significant differences in the use of force by officers or citizen complaints against officers whether police were wearing BWCs or not (Yokum, Ravishakar, & Coppock, 2017). The premature adoption of BWCs, at least in our nation’s capital, resulted in the purchase of
costly equipment and video storage space that did not meaningfully contribute to reducing violence on the part of officers or the complaints that citizens made about unnecessary force used in law enforcement.

We have been inclined to adopt quick fixes in the field of speech language pathology as well. Many may recall that years ago, Scientific Learning Company began marketing a program called Fast ForWord-Language (FFWD-L), claiming that it served as “glasses for the ears” (De Anda, 2000) and led to remarkable gains in language skills in children with language impairments (Tallal, Miller, Bedi, Byma, Wang, Srikantan, Nagarajan, Schreiner, Jenkins, & Merzenich, 1996). FFWD was touted as a cure for language impairments in scientific articles (Merzenich et al., 1996; Tallal et al., 1996) and in the popular press, which led to its widespread use across the U.S. and the world (Kamhi, 2004). The basic idea was that FFWD-L would improve children’s temporal auditory processing abilities which in turn would lead to improvements in their general language skills (Scientific Learning Corporation, 1998). The program was based on basic research on neural plasticity, and the connection between auditory processing abilities and language skills seemed reasonable. Therefore, use of the program gained momentum very quickly. We know now that FFWD-L is no better than many other language intervention approaches, but it took years of rigorous scientific research to overcome the claims of a quick fix (Cohen, et al., 2005; Fey, Finestack, Gajewski, Popescu & Lewine, 2010; Gillam, et al., 2008; Strong, Torgerson, Torgerson, & Hulm, 2011). Unfortunately, as this example illustrates, there are not always quick solutions to complex problems no matter how straightforward they seem.

Today, we find ourselves in a similar set of circumstances surrounding working memory (WM) capacity training. WM has been defined as the storage and processing of information for
use in completing cognitive tasks (Baddeley, 2003; Cowan, this issue). Speech Language
Pathologists (SLPs) are very interested in the potential benefits of working memory training
because their students with Developmental Language Disorders (DLD) often perform more
poorly than their peers on a variety of WM tasks. For example, the difficulties that students with
DLD have with working memory has been linked to poor vocabulary (Montgomery, Magimairaj,
& Finney, 2010) phonological awareness (Alloway, Gathercole, Adams, & Willis, Eaglen &
Lamont, 2005) sentence comprehension (Montgomery, Gillam, & Evans, in press), literacy
(Swanson, Zheng, & Jerman, 2009; Wang & Gathercole, 2013; Werfel & Krimm, 2017), and
mathematical calculation and problem solving (Alt, Arizmendi, & Beal, 2014; Cowan, Donlan,
Newton, & Lloyd, 2005; Fazio, 1999; Nys, Content, & Leybaert, 2013; Peng, Namkung, Barnes,
& Sun, 2016).

The fact is, working memory is one of the best predictors of performance on tasks
measuring a variety of linguistic and academic abilities (Cowan, et al., this issue; Peng,
Namkung, Barnes, & Sun, 2016; Swanson, 1994). So, it stands to reason that if we could
improve WM capacity, we may also be able to improve performance on other language and
academic tasks. Unfortunately, research into the transfer effects of WM capacity interventions to
linguistic and academic abilities has not yielded sufficiently compelling results.

Current State of the Evidence

A large number of studies have been conducted to investigate whether WM capacity
training is associated with gains in WM capacity along with a wide range of cognitive and
academic skills such as fluid intelligence, decoding, reading comprehension and math skills. A
number of systematic and meta-analytic reviews have been conducted to critically examine the
findings of some of these well designed studies (Au, et al., 2015; Karbach & Verhaeghen, 2014;
Melby-Lervåg & Hulme, 2013; Melby-Lervåg & Hulme, 2016; Melby-Lervåg, Redick, & Hulme, 2016; Randall & Tyldesley, 2016; Sala & Gobet, 2017; Schwaighofer, Fische, Bühner, 2015). Details of the meta-analytic reviews are displayed in Table 1. One consistent finding across all of these reviews was that WM training resulted in immediate improvement in tasks that were trained, or those similar to those that were trained (near transfer effects). Unfortunately, WM training only occasionally has been linked to improvements in skills, like language and reading, that are thought to be related to WM but are unlike the tasks that were directly trained (far transfer effects).

Melby-Lervåg and Hulme (2013) conducted a meta-analysis to determine whether working memory programs were effective in bringing about improvements in working memory capacity as well as other cognitive and academic abilities. Their review included 23 studies with 30 group comparisons made up of clinical and typically developing children and adults. They reported near transfer effects of WM training to verbal WM tasks however, these were not maintained 9-months post-treatment. Near transfer effects were also found for visuospatial WM tasks that demonstrated some degree of maintenance over time. In terms of far transfer, WM training did not appear to generalize to other skills including nonverbal and verbal ability, inhibitory processes in attention, decoding or math. Their conclusion was that memory training does appear to result in short term, specific training effects to working memory tasks but not to other cognitive or academic skills in typically developing children and adults. These results were replicated in a follow-up meta-analysis by these same authors in 2016 (Melby-Lervåg et al., 2016).

Karbach and Verhaeghen (2014) conducted a meta-analysis to investigate whether process based training protocols that target general processing capacities (e.g., processing speed,
executive functioning) or working memory resulted in improved cognitive functioning (e.g., fluid intelligence) in healthy, elderly persons. A total of 49 studies with 61 different experiments or independent subject groups were included. The near transfer measures included tasks that were not directly trained but measured the same construct as the training task. The far transfer measures included tasks that measured a different construct than the one trained. They reported significant near (Hedges g = .47) and far transfer effects ($d^+ = .37$; Hedges & Olkin, 1985) for the studies they included. Interestingly, Melby-Lervåg and Hulme (2016) re-analyzed the data related to far transfer that was reported by Karbach and Verhaeghen (2014) and found somewhat smaller, statistically non-significant effects (Hedges g = .21). Their results differed because Melby-Lervåg and Hulme (2016) corrected for weaknesses in Karbach and Verhaeghen (2014) including non-transparent meta-analytic methods and reporting, inaccurate effect size calculation, and lack of separation of results of studies with active and passive control groups. Their findings suggested that far transfer effects reported earlier may not be as robust as had previously been reported.

Another recent meta-analysis reported small statistically significant effects for improvements in far transfer to fluid intelligence measures after working memory training with n-back tasks (Au, Sheehan, Tsai, Duncan, Buschkuehl, & Jaeggi, 2015). This meta-analysis contained 20 studies that included 30 different treatment groups and 24 different control groups comprised of healthy adults between the ages of 18 and 50. All of the studies included in the meta-analysis trained participants on some form of adaptive n-back task, included a control group, and used some form of fluid intelligence (Gf) assessment as the outcome measure. They reported a small significant net effect of n-back training on Gf outcome measures ($g = 0.24$, SE = .07). Melby-Lervåg & Hulme (2016) repeated the analysis with the same studies, but corrected
for effect size calculation errors with different results (Hedges g = 0.13). They authors also expanded the search to include more data-bases and re-analyzed the data and found another small but significant effect size (Hedges g = .10). Their findings suggested that the results from the Au et al., study may not reflect the true nature of the evidence that currently exists.

Schwaighofer, et al., (2015) conducted a meta-analysis to investigate near and far transfer effects of working memory training to a number of visual, verbal and nonverbal tasks as well as to decoding and math skills. They included 47 studies with 65 group comparisons. Participants ranged in age from 4-71 years old. The experimental instructional programs included any adaptive computerized program that targeted verbal and/or visuospatial working memory skills. The results of the study showed that there were immediate and sustained (8 months follow up) near transfer effects for verbal and visuospatial STM, as well as verbal and visuospatial WM. Schwaighover et al. found small immediate far transfer effects from WM training to nonverbal and verbal abilities but these improvements were not sustained over time. They found no evidence of far transfer effects from WM training to decoding or mathematical abilities.

Randall and Tyldesley (2016) conducted a systematic review that included 8 studies investigating near and far transfer effects of three programs designed to improve working memory capacity (CogMed, Robomemo, Memory Booster) for children ages 5-12 with and without learning differences. They reported near transfer effects ranging from Cohen’s d = .28 to 1.19 for verbal and visuospatial short term memory and working memory tasks. They reported mixed results of far transfer effects. There were no reported far transfer effects to IQ or ADHD symptoms and variable results regarding literacy and numeracy. Only one of the five studies reporting literacy outcomes reported significant far transfer effects in this area (Cohen’s d = 0.88 and 0.91; Dahlin, 2011). While the effect sizes reported are large, they should be interpreted
with caution because of methodological concerns (i.e. effect size calculation methods and lack of equivalency at baseline). Results for WM training transfer to numeracy were also inconsistent with effect sizes ranging from non-significant to Cohen’s $d = 1.15$. It is evident from the results of this synthesis that the true effects of WM training on the near and far transfer skills of children with and without learning differences are as yet unclear. Further investigation is needed in order to determine the efficacy of these programs with children.

Recently, Sala and Gobet (2017) conducted a meta-analysis to examine near and far transfer effects of working memory training in typically developing participants ages 3 to 16. They reported that, like many meta-analyses we have already discussed, participants in the studies they reviewed experienced significant near transfer of WM training to WM related skills (Hedges’ $g = .46$) such as verbal and visuospatial working memory. There were no far transfer effects to measures of fluid intelligence, cognitive control or academic abilities such as literacy or science, but they did report a modest effect size for math related skills (Hedges’ $g = .20$). The authors commented in their article that they had limited confidence in the findings reporting math skills due to methodological concerns.

Two relevant studies were not included in Sala and Gobet (2017) because they involved children with DLD and children who were at risk for learning disabilities and were published after the meta-analysis search was completed. Because they are directly related to children SLPs serve, they bear examination here. Vugs, Knoors, Cuperus, Hendriks, and Verhoeven (2017) conducted a pilot study with 10 children with DLD aged 8 to 12 years to explore the use of a computerized executive functioning training program with children with DLD. The program addressed three areas of executive functioning: visuospatial working memory, inhibition, and mental flexibility. Children’s were assessed at pretest, posttest, and after 6 months on trained
tasks and far transfer measures including verbal working memory, attention, planning, and verbal fluency at immediate post-test and 6 months after intervention. Additionally, behavioral ratings from parents and teachers were collected. Children participated in 6 weeks of intervention total. Children improved significantly on cognitive flexibility immediately after intervention and showed significant improvement on all three trained tasks at 6 months post-intervention. They also showed significant improvement in two attention measures and behavioral ratings from parents and teachers at 6 months post-intervention.

While these results may appear promising at first glance, a number of issues, which the authors also noted, indicated that this intervention requires further investigation before strong conclusions about its effect on WM and transfer measures are made. The most concerning issue with this study was the lack of control group. Without a control group, validity confounds such as history and maturation cannot be ruled out as explaining the results (Shadish, Cook, & Campbell, 2002). This is particularly true since the majority of the significant effects were found at 6 months post-intervention rather than at immediate posttest. Even if this study had included a control group and obtained the same results, it is interesting to note that neither visuospatial WM, which was trained, nor verbal WM, which was not trained, improved at immediate posttest. It was only at 6 months post intervention that visuospatial memory measures showed significant improvement. Neither verbal WM nor verbal fluency showed significant improvement, suggesting that this intervention, as presented, would not be optimal to address the core deficits of children with DLD.

The second article published after the Sala and Gobet (2017) meta-analysis was Peng and Fuchs (2017). These authors conducted an RCT of young at risk children’s performance on WM and reading comprehension tasks after WM intervention. Fifty-eight children in 1st grade who
were at risk for learning difficulties were randomly assigned to 3 groups: a WM with rehearsal strategy group, a no-rehearsal WM intervention group, and a control group. The intervention groups participated in one-on-one non-computerized WM training. The rehearsal group was trained to use an immediate rehearsal strategy coupled with the WM task. When results were corrected for multiple comparisons, no significant group differences emerged. Without correction, the rehearsal strategy instruction group showed the most improvement, which would suggest that if there were true differences between the groups, explicit instruction in the use of a strategy was a powerful ingredient in the intervention. This is consistent with previous research on strategy use with children at risk for learning disabilities (Dole, Brown, & Trathen, 1996). Further, intervention was delivered in person rather than solely by computer, which may have resulted in greater improvements, consistent with Schwaighofer et al.’s (2015) finding that supervised training resulted in larger effects than unsupervised training.

Given that the findings related to far transfer effects of basic memory training are, at best, equivocal, it appears that researchers have yet to demonstrate the feasibility of WM training for influencing functional abilities in language, reading, or mathematics. Such outcomes are insufficient to support widespread implementation of WM capacity treatments as a means to improve these related skills in environments where evidence-based practices are valued. Clinicians who decide to provide children with basic training on WM tasks should be sure to measure both near and far transfer skills. They should be ready to pull the Andon rope if their functional language and academic measures (far transfer) do not change in a reasonable amount of time. Better yet, clinicians might consider focusing their interventions on those communication skills that are problematic and that relate to academic success, and worry less about underlying cognitive skills (like working memory) that may or may not relate to functional
language and academics. Said a little differently, the cognitive and academic interventions we use should incorporate the functional skills we mean to train (Melby-Lervåg & Hulme, 2013; Sala & Gobet 2017).

**Development of WM and other cognitive processes**

If we are to determine how best to use WM interventions with children with DLD, if at all, we must understand how WM and other cognitive processes develop. We know that performance on simple memory span tasks and working memory measures improves as individuals age (Gathercole et al., 2004). There is an assumption that individuals have a limited number of items that they can recall immediately, referred to as memory capacity. However, the number of items that one is able to recall may not be equal to the number of “spaces or slots” in their capacity (Cowan, 2016). This is because a complex item might be converted to more than one “chunk,” whereas multiple, potentially related items might comprise a single chunk. Research has shown that infants are capable of chunking even before they have developed language skills (Feigenson & Halberda, 2008). Therefore, it is possible that in addition to capacity increases (if capacity increases at all), there may be developmental change that is a result of improved chunking (or processing) efficiency. Gilchrist, Cowan and Naveh-Benjamen (2009) proposed a method for calculating how many sentences or chunks could be remembered “mostly intact.” The resulting outcome measure was called chunk access, which was calculated by totaling the number of sentences in a paragraph or story in which at least one content word (noun or verb) was recalled. In their study, the ability to chunk information improved over time for children ages 7 to 12; and in later studies was shown not to be solely mediated by increases in knowledge (Cowan, Ricker, et al., 2015).
A related concept involves the development of attention and the ability to retrieve information from long term memory. Cowan (2016) proposed an embedded process model to explain how attention and memory may interact. The model suggests that incoming stimuli from the environment automatically activates attributes (words, phrases, sentences) stored in LTM that are related to it. After activation, these attributes are subject to decay and to interference from other items in long term memory that may or may not be relevant. However, items that were in the focus of attention may remain partially active for a time after they are no longer the central focus of attention. Items can get into the focus of attention through automatic or deliberate routes. The automatic route to action involves noticing that incoming information is different in some way making it conspicuous. Deliberate routes to action are governed by central executive processes that determine what should be acted upon. When automatic and deliberate routes from attention to action do not agree, the deliberate route may need to override an automatic, incorrect response. It is possible that through the automatic route, individuals may experience a deja vu effect that makes the wrong answer compelling and difficult to suppress. At that point, in order to select the correct response, executive functioning processes should engage the deliberate route to action to select the correct response over the recently activated, incorrect response.

It remains to be determined if the changes we see in WM over time are the result of an increasing number of “slots” that may be filled (increases in capacity) or in the ability to adequately moderate executive functioning to deliberately control the focus of attention to retrieve accurate information from LTM (increases in processing efficiency). Given the dynamic interaction between factors related to memory and other cognitive skills, it is unlikely that training in only one aspect of the system, such as WM capacity or the ability to shift attention quickly, can affect large changes in the other mechanisms that are involved. This is particularly
relevant for children with DLD who demonstrate limitations in WM capacity as well as other cognitive and strategic processing skills.

**WM and children with DLD**

It is well established that children with DLD perform more poorly on working memory tasks than their age-matched, typically developing (TD) peers, and in many cases, even their language-matched TD peers (Archibald, 2016; Claessen, Leitão, Kane, & Williams, 2013; Marton & Schwartz, 2003; Marton & Eichorn, 2014; Montgomery et al., 2010; Weismer et al., 1999; Wong, Ho, Au, McBride, Ng, Yip, & Lam, 2017). Research consistently shows that children with DLD differ in how much linguistic information they can store, and also in their development and use of other cognitive skills and strategies related to WM efficiency.

One such example relates to the use of attentional resources. It has been suggested that children with DLD have difficulty allocating attention to the task of loading working memory with relevant versus irrelevant items (Cowan, 2010; Kane, Bleckley, Conway, & Engle, 2001; Vogel, McCollough, & Machizawa, 2005). For example, Morton and Eichorn (2014) tested this hypothesis in three groups of children (15 children with SLI, 15 age-matched peers, and 15 language-matched controls) using listening span and nonword repetition tasks. In the listening span tasks, children were asked to listen to a sentence, recall the last word, and answer a question about it. These sentences (containing their words) were then presented to them as a listening span task. Children were also asked to repeat nonwords that contained meaningful and nonmeaningful syllables. Children with SLI performed more poorly than both groups on both of the nonword tasks. Children with SLI performed more poorly on comprehension tasks, and experienced significantly more interference errors in recalling words during the listening span tasks than their age and language matched peers. These findings were taken as evidence that children with SLI
may not use efficient retrieval strategies to recall information to support WM performance, partly because they have difficulty suppressing irrelevant material. It is possible that students with DLD retrieve more irrelevant items because they are more susceptible to interference from competing information that has recently been activated and/or is stored in LTM (Marton & Eichorn, 2014).

Studies of discourse processing have shown that children with DLD may also have difficulty generating inferences during authentic comprehension tasks because of problems with WM, attention, and difficulty retrieving information from LTM. Laing and Kamhi, (2002) investigated discourse comprehension in 3rd grade children with and without DLD who participated in think aloud tasks involving narrative passages. Children were asked to listen to stories, one sentence at a time. After hearing each sentence, children were asked, “What do you know about the story now?” Their responses were coded as to whether they were an inference, a paraphrase, or a repetition of the information contained in the passage. Then, each statement or “verbal protocol” was analyzed for accuracy. After hearing the entire story, children were asked to recall the passage and to answer comprehension questions about it.

The children with DLD demonstrated significantly poorer WM than their typically developing peers, answered fewer comprehension questions accurately, recalled less information from the stories, and made more errors in generating accurate inferences than their typically developing peers. The researchers did not conduct a formal evaluation of interference errors; however, it was observed that children with DLD reported information in their story recalls that was not in the original passage. When these statements were examined, it was noted that the inaccurate statements were often part of the student’s think aloud. For example, a student reported during the think aloud task that the character, Bill, could see now, because he had gotten glasses from the doctor. During the story recall task, the student recalled that the character was
no longer blind because he was now wearing glasses. In the passage, glasses were never mentioned, and in fact, Bill’s blindness was cured by an encounter with a magic lake of water.

Gillam, Fargo and Robertson (2009) found similar results for 4<sup>th</sup> grade children as they participated in think aloud tasks involving comprehension of expository passages. The participants with DLD demonstrated poorer WM than children developing typically; were less likely to recall information from paragraphs accurately, and were twice as likely as their typical peers to include irrelevant or inaccurate information in their story summaries.

Together, these findings provide some support for Marton and Eichorn’s suggestion that children with DLD have difficulty accessing information from LTM to support their WM performance. Marton and Eichorn recommended, and we concur, that students with SLI may require direct strategy training to improve their resistance to interference and their ability to retrieve information from LTM in support of WM efficiency.

**Strategy instruction in support of WM**

A number of strategies and functional activities may help improve working memory efficiency as well as to improve the proficiency with which students allocate their attentional resources while they perform complex tasks. The following discussion is based on selected studies that do not represent an exhaustive review of the subject, but serve to highlight techniques which have been shown to be efficacious in supporting children’s working memory and attentional functioning. It is likely that these strategies and skills will be useful regardless of whether a practitioner ascribes to interference-related or decay-related explanations of working memory limitations. Careful progress monitoring of the outcomes of such interventions are encouraged.
One such strategy is to teach children an organizational framework. Organizational frameworks may improve WM efficiency and the ability to retrieve information from LTM because they impose a structure upon information that is coming in. This may allow individuals to focus on relevant information while providing them with a place to “hang the information” while they search LTM. Kintsch proposed a theoretical framework for understanding how all this works called the Construction-Integration Model (CI; Kintsch, 2013). According to the CI model, comprehension requires that we construct a mental representation of what we read or hear (textbase) while scanning knowledge in LTM to determine what it means to us. Kintsch proposed that comprehension requires that we construct an overall model of the information using hierarchical relationships among key ideas (macrostructure). Integration involves the formation of a situation model which is formed by comparing background knowledge and experiences to the textbase that has been constructed. Finally, the new information in the textbase is strategically and consciously linked to old information and stored in long-term memory.

The use of icons, visual organizers, concept/knowledge maps, mental imagery and graphic organizers have a strong evidence base for improving listening, higher level language skills (ex. inferencing), writing, grammar, vocabulary/concept learning, reading and math (Dexter & Hughes, 2011; Joffe, Cain, & Maric, 2007; Kansizoglu, 2017; Nesbit & Adesope, 2006). Most studies have included children who are developing typically, are low readers, or are at risk, however a number of experimental and quasi-experimental studies have shown that teaching strategies to aid storage and retrieval of information in long term memory may also be effective for students who have language learning problems (Gillam, Hartzheim, Studenka,

In a compelling investigation of the power of the use of icons in aiding memory for social studies content, Nathanson, et al., (2007) asked whether teaching story structure using cue cards would retrospectively improve recall of a Mexican history lesson that students had been taught 2 weeks earlier. Two weeks after having the history lesson, 39 children with learning disabilities in first through fifth grades were randomly assigned to a treatment group who received instruction in story structure (narrative elaboration treatment; NET) or to a control group that did not receive the instruction. Students in the NET group were taught to associate cue cards with aspects of character and setting details, actions, and feelings. Students in both groups viewed videos and were asked to recall what they had seen. Students in the NET group were taught to use the cue cards to organize their retellings while students in the control group were not. Findings revealed that participants in the NET group remembered significantly more information from the history lesson than children in the control did. These findings speak to the powerful effects of using a strategy to improve chunking information for retrieval from LTM. The students were able to retrieve the information they had stored about the lesson from their long term memories by using an organizational framework to help them decide what they needed to remember and then to retrieve it from LTM for recall.

The components of the C-I model informed each of the lessons in a narrative program used by Gillam et al., (2015) who conducted a study to examine whether this program was associated with improving students abilities to store and retrieve discourse level information. Participants in this multiple-baseline, single-subject design study were 3 children with Autism Spectrum Disorders (ASD) who had higher language skills (CELF-4 Standard Scores of 79, 85,
114) and 2 children with ASD who had lower language skills (CELF-4 Standard scores = 62, 48). Data were collected for lagged baseline and intervention phases over a 3-month period. The narrative intervention taught students icons, graphic organizers and associated verbal cues (e.g., “all of a sudden” for initiating event; “they decided to ___” for plan; “uh oh” for complication “in the end” for consequence) to help them remember critical story elements. Students were also taught ways to represent causal connections between elements by using words like “because” and “so.” Children were asked to create their own stories from single scene picture prompts at each baseline and every other treatment session. The children made significant gains in telling more complete and complex stories over the course of intervention. Gillam et al., (2015) believe that students were using the organizational framework that was taught to them to organize information in LTM to create novel stories. However, the retell data were more compelling for making this point.

As part of the study, children were also asked to retell stories after every baseline session and every other intervention session. The stimulus stories ranged in length from 18-22 utterances and contained between 200 and 300 words. All of the stories contained 2-3 episodes (initiating event, action, consequence) and one or more complicating actions (something that gets in the way of the character’s goal-motivated actions). The examiners read each story to the student and then asked the student to tell the stories back to them. No pictures, prompts, icons or graphic organizers were provided for students during the retell sessions.

The Monitoring Indicators of Scholarly Language (MISL) rubric was used to measure narrative proficiency (Gillam, Gillam, Olszewski, Fargo & Segura, 2017). There are 7 items on the MISL that were used to measure macrostructure elements and 6 items that measured microstructure elements. The microstructure scores are not relevant here and will not be
discussed. For macrostructure, each item (character, setting, initiating event, internal response, plan, attempt and consequence) was scored as absent (score of 0), emerging (score of 1), present (score of 2) or elaborated (score of 3). Each story retell was also coded as to whether it contained 2 or more initiating events, 2 or more actions related to those initiating events, a complicating action, and a clear consequence that tied it all together. In order to meet the criteria for recalling a complex episode the student had to earn a score of 2 or higher in each of the categories for initiating event, action, and consequence recorded on the MISL.

The data for the stories at baseline and for last 4 elicitation probes during intervention were examined in order to judge whether knowledge of the framework that was taught was associated with greater recall of information from stories. Of the three children with ASD who had higher language skills, one recalled 40% of complex episodes contained in retells at baseline and 100% during the last four data collection points, one recalled 33% of the complex episodes in stories she heard at baseline and 75% during the last 4 data points, and the third recalled 11% of the complex episodes during stories at baseline and 100% during the last four data collection points. Of the two children with ASD who had lower language skills, one recalled no complex episodes during baseline but 75% during the last 4 data collection points. The other student who had the lowest language skills was unable to recall any complex episodes during baseline or intervention. These findings suggested that narrative language intervention that provided students with ASD with a clear organizational structure for stories led to much higher memory for the critical elements in stories that the children heard.

According to C-I theory, the reader (or listener) must hold information in WM long enough to scan through LTM and connect incoming information with what is already known. The listener is continually “making sense” of what is coming in by comparing it to what is
known. In the case of a student who has poor knowledge of the topic, more structure is needed in order for him or her to know “what” to hold onto in the incoming information while at the same time searching for what is “relevant” in LTM. In the case of our narrative intervention, we taught children a structure or framework to use while constructing an elicited story or in recalling a story that they heard. The framework consisted of selected story elements (character, setting, take off [initiating event], internal response, plan, action, consequence and wrap up [summary with feelings]) and the causal relationships between them marked by because, and so.

The hierarchical framework that was taught included selected story elements (character, setting, initiating event, internal response, plan, action, consequence) and the causal relationships between them. The causal framework was taught through modeling, repetition and practice. An example of how this framework is taught follows.

Clinician/Examiner: Bob and Max, they are the characters, were walking in the forest, during the day, to get to their campout. Forest, daytime and campout all represent setting elements. Two are places, forest, campout, and one is a time of day (during the day). All of a sudden, those are our cue words for take-off, a bear showed up in their path. That’s how the story gets going. That is the take off. The bear shows up in the pathway, and they were afraid. Afraid is a feeling. Why do you think they are afraid.

Child: Because the bear might eat them.

Clinician/Examiner: Yes, they are afraid because (asking why to get at the causal framework) the bear might eat them. So (to highlight causal framework; relationship between take off, feeling and plan) they decided, that’s the plan, that’s what they thought in their head that they would do; they decided to run away and hide from the bear in a tree. So, (to highlight relationship between take off, internal response, plan and action),
they ran and climbed the tree. But, ah, oh (cue words to for complication), the bear fell asleep at the bottom of the tree. Now they can’t get down and go to their campout. That’s a complication. When the bear fell asleep at the bottom of the tree, it got in the way of their action. Remember, they planned to get away from the bear by running and climbing up the tree, and that’s what they did. But the bear is still there so they still can’t go to their campout. After a while, the bear woke up and left, so Bob and Max climbed down from the tree and went to their campout. That’s the landing, that’s how the story ends.

They escaped the bear and made it to their campout, and they were very happy that they were eating marshmallows and not getting eaten that day by a bear. That the wrap up. The wrap up is what happened in the story and how the characters felt about it.

As students learn the framework, they may be better able to moderate attention, making the most of the working memory capacity they have available to complete complex tasks such as recalling aspects of stories. In the following example, note how the student overtly marked aspects of the organizational framework, in this case, macrostructure, in order to remember parts of the story.

Original Story: One spring afternoon, Eli and his sister Sara ran off to play baseball at Logan Park. Flowers were blooming and the weather was warm and sunny.

“Throw the ball,” yelled Eli. Sara threw the ball as hard as she could. Eli jumped high into the air to catch the baseball. The ball flew so high, he lost it in the tree branches above him. Before he knew what happened, an egg, not a ball, was falling out of the tree. Surprised, he reached out his baseball glove and carefully caught a tiny, blue robin’s egg. Cara ran over and looked in Eli’s glove. “Oh no,” she cried. “What do we do? Is it broken?” “No,” said Eli, relieved. “The egg is safe.” They decided to try and put the egg
back into the nest so it would have a chance to hatch into a baby robin. Eli carefully handed the egg to Sara and said, “Climb up on my shoulders and put the egg back into the nest.” Sara wanted to be extra careful with the egg because she did not want to break it. Sara held the egg gently in one hand and climbed up onto Eli’s shoulders with her other hand. She placed the egg back into the nest and jumped down onto the grass.

“Whew,” she said, “That was so close.” “I’m so glad you were careful or else the little robin’s egg would have broken!” “Me too,” said Eli. “You can really throw the ball hard. But next time, let’s be sure to stay away from the trees. Maybe I’ll have a chance at catching a ball, not an egg!”

The story that follows is the retell told by a student early in intervention, when he was just beginning to learn the organizational framework. The aspects of narrative structure that we taught are highlighted in bold.

So once upon a time there were two **characters** named Eli (and and) and his sister named (uh) Abby. (And um) and they were outside in a sunny afternoon [*included setting elements of “where” and “time”*]. (And so) but all of the sudden when Elijah threw a ball into a tree (it) it almost broke a robin’s egg. **So they decided** to put the egg back into the nest. And Elijah said, “Whew that was close.” And they jumped back. Elijah jumped off Abby’s shoulders onto the ground. (And, but) **so in the end** they would remember that next time that they throw a ball that it doesn't get into the tree/s.

In his retell, this student, who could not recall even simple episodes during baseline, appeared to utilize the framework we taught to help him organize his knowledge and thus remember more of the narrative. We believe that the basic structural elements and causal connections we taught were automatically activated from LTM as his focus of attention was on
the incoming narrative. This structure may have enabled him to attend to the critical elements of
the narrative and to suppress competing information which may also have been activated from
LTM as he listened. Then he integrated the basic structure we taught with the relevant
information from the narrative to produce the final retell which was relatively well organized and
free from irrelevant information.

Other organizational frameworks have also been used to support storage and retrieval of
higher level language skills such as inferencing. For example, Murza and colleagues (2014)
investigated the use of graphic organizers, advanced organizers, cue cards, structured notes, and
post organizers to support memory performance in service of accurate inference generation
during reading for 25 adults with high functioning Autism Spectrum Disorder (HF-ASD). Their
randomized controlled trial (RCT) showed that teaching adults with HF-ASD to use a four-step
reading comprehension strategy (Ask yourself a question, Consider the text, Think about what
you know and take a good guess, Check your Guess) was effective in improving inference
generation during reading.

More recently, Barth & Elleman (2017) conducted an RCT to examine the impact of
teaching inference strategies using graphic organizers to 66 middle school students identified as
struggling readers. Students in the treatment group were taught 1) clarification using text clues,
2) understanding character perspectives and author’s purpose, 3) activating and using prior
knowledge, and 4) answering inferential questions. The intervention was short, only 10 days, but
yielded large effect sizes for the content that was taught (Hedges’ g = 1.37) and moderate effect
sizes on a standardized measure of reading comprehension (Hedges’ g = .46). These studies
highlight the importance of providing an organizational framework for improving the efficiency
of WM as well as other complex tasks.
Can we improve WM using more functional tasks?

There is some research that has shown that we can improve functional skills directly, and in turn improve WM efficiency. For example, Van Kleeck, Gillam & Hoffman (2006) showed that improvements in WM may be brought about by training phonological awareness skills. They recruited 16 preschool children with DLD who received intervention in groups targeting PA skills, 15 minutes twice weekly for 30 weeks. Children were also given nonword and word span tasks (single syllable, multiple syllable). Following intervention, the children’s PA skills improved significantly. However, these same children also made significant improvements on the memory measures used.

Similarly, Park, Ritter, Lombardino, Wiseheart, and Sherman (2014) examined the impact of PA training on improving PA, decoding and verbal WM abilities in children with DLD. A total of 50 school-age children with DLD and word reading deficits were randomly assigned to either an experimental or control group. Both groups received individual language intervention targeting narrative skills for four, 1 hour sessions for 4 weeks. The experimental group also received an additional 20 minutes of phonological awareness training each session. The children in the experimental group, who received PA training, outperformed the children in the language-only group on measures of PA and word decoding. They also improved more than the non-PA control group on measures of verbal WM such as word list recall. The findings suggested that school-age children with DLD and word decoding deficits may make significant gains in untrained verbal WM skills by improving their PA skills.

Far transfer generalization of training in narrative intervention and phonological awareness intervention to WM provides proof of concept evidence that working on functional tasks that are closely related to academic abilities, like decoding and reading comprehension,
may also affect more basic working memory skills. Given the equivocal results of far transfer effects of direct memory training for many different kinds of children at different ages, clinicians should be ready to pull the Andon rope and provide children with functional language intervention that is designed to improve the academically-related communication skills that children are lacking while the research continues to examine how best to improve working on memory capacity directly.

**Other strategies to improve WM efficiency**

Other factors related to WM task performance that may affect the ability of children with DLD to benefit from intervention are important for SLPs to consider when designing treatment plans for children with DLD. One such factor is how lowered self-esteem and repeated failure can affect performance on WM tasks. Autin and Croizet (2012) conducted a fascinating set of studies to investigate the extent to which WM efficiency could be improved by alleviating concerns that participants had about being incompetent. There is a line of research that has shown that when participants feel they are going to look like they are not intelligent, they perform more poorly on the complex tasks they are being asked to do. Reportedly, negative thoughts and feelings of incompetence result in “off task” thoughts that tax available cognitive resources. These off task thoughts are said to compete for attentional control, which in turn steals from the already limited capacity available in WM, leading to poor performance (DeCaro, Thomas, Albert, & Beilock, 2011). A strategy that has been used to minimize off task thoughts related to lowered self-esteem is termed “difficulty with reframing” and was tested across a series of studies conducted by Autin and Croizet (2012).

In their first experiment, Autin and Croizet randomly assigned 111, 6th graders to one of 3 conditions. In the first condition, difficulty with reframing, students completed 3 series of 2
difficult anagrams that were unsolvable in the time frame given to them, thus ensuring failure. They were told in a debriefing fashion, that when they experienced processing difficulty it was the normal result of learning. Then they were given a listening span task. Students in the second condition, difficulty without reframing, performed the same anagram tasks but were given no debriefing. Then, they participated in listening span testing. Students in the third condition, the standard condition, did not complete the anagram tasks, only the listening span task. Findings revealed that the students in the difficulty with reframing condition performed better on the listening span task than children in the difficulty without reframing condition and the standard condition.

In their second experiment, the impact of this strategy was explored in terms of its impact on reading comprehension, a task that relies heavily on WM. One hundred thirty-one 6th graders participated. Rather than a listening span task, all students were given a difficult reading comprehension task to complete. As in the first experiment, prior to the comprehension task, one group completed unsolvable anagrams and received difficulty with reframing debriefing; the second group did not receive reframing, and the third group completed the comprehension task only. A fourth group also participated in this study. For this condition, students were able to successfully solve the anagrams they were given. They did not receive reframing because the task was not difficult. Then, they participated in the reading comprehension task.

A similar pattern of findings emerged. Students who learned that having difficulty was a normal aspect of learning new information performed significantly higher on the reading comprehension measure than students in the other conditions, including the one in which students were successful at solving the anagrams. Together, these findings support the notion that when students are confronted with a difficult task, their own feelings of lowered self-image
can derail their ability to efficiently store and process information. “Difficulty with reframing” is a means to improve WM efficiency, rather than capacity, by reducing an emotional threat that may be induced when students must undertake very demanding tasks.

This information has very important implications for our students with language disorders who have been shown in a number of studies to experience lowered self-esteem compared to their peers who were developing typically (Durkin, Toseeb, Botting, Pickles, & Conti-Ramsden, 2017; Jerome, Fujiki, Brinton, & James, 2002; Wadman, Durkin & Conti-Ramsden, 2008). Furthermore, issues of lowered self-esteem may have a greater impact on older students than younger ones. For example, Jerome et al., (2002) examined self-esteem in 46 children between the ages of 6 and 9 and 34 children between the ages of 10 and 13. Some were developing typically and others were diagnosed with DLD. There were no differences in self-perception scores for the younger children. However, older children with DLD perceived themselves more negatively on measures of academic competence, social acceptance and behavioral conduct than their peers who were developing typically.

This finding was confirmed more recently by Durkin and colleagues (2017) who reported data for over 200 participants with and without a history of language impairment between the ages of 17-24. The participants completed measures of language, nonverbal IQ, shyness, global self-esteem and self-efficacy. The participants with a history of language impairment scored lower than age-matched peers on self-esteem, shyness and self-efficacy measures. All of the differences were associated with moderate to large effect sizes.

It is possible that “difficulty with reframing” techniques may be very valuable in helping students to direct their already limited resources toward the task they are trying to complete, rather than using them to manage emotional feelings of inadequacy. Other interventions that
may benefit students with language impairments may include teaching them self-management strategies, strategies for dealing with social situations, helping them to modify and/or manage their negative expectations and to increase their confidence in their abilities to learn new skills (Durkin et al., 2017).

**Summary**

Fifteen to twenty years ago, Alan Kamhi, wrote a series of articles exploring factors that contributed to the kinds of intervention approaches clinicians choose to use in the treatment of childhood speech and language disorders (Kamhi, 1999; 2004; 2006). In those papers, Kamhi talked about the fact that clinicians sometimes had to make decisions to use approaches that lacked scientific evidence because there were so many procedures that did not have “high levels of evidence” to support their use. He suggested that the driving reason to use such techniques was that they appeared to be effective in bringing about change in critical, functional behaviors. However, he also stated that changes in behavior should be experimentally validated, meaning that clinicians should be able to show that it was the treatment, rather than other variables, that brought about said changes.

We explained earlier in this paper that the evidence of cognitive, academic, and linguistic improvements following WM intervention is equivocal. The premise that WM interventions are effective for children with DLD rests partly on the assumption that the factors that limit WM capacity in children with DLD are the same as those in other populations. Research has shown that this may not be entirely true. Recall that Marton and Eichorn (2014) and others (Laing & Kamhi, 2002; Gillam, et al., 2009; Gillam et al., 2015) showed that children with DLD demonstrate specific weaknesses in using what they already know (in LTM) to support WM performance, partly due to an inability to suppress information that is not relevant or accurate for
the task and partly due to not knowing what to pay attention to. Until there is more clarity on how WM is limited for children with DLD, and how those limitations interact with LTM representations, it may be premature to think that currently available treatments aimed at increasing WM capacity are as targeted as they ultimately should be. For these reasons, we cannot yet draw consensus conclusions on the usefulness of WM treatments for children with DLD.

You may recall earlier in the paper we discussed how pulling the Andon rope, while tedious and time-consuming, led to a long reign of quality in the automobile industry for the Toyota Company. Unfortunately, Toyota stopped this process in the 90’s in their race to be number one in the industry which led to the recall of over 10 million vehicles and seriously hurt their reputation. While we are not suggesting that WM interventions have no place in our tool kit for children with DLD, we do encourage clinicians who are using direct WM treatments to look very carefully at client data and to pull the Andon rope when they are not seeing improvements in the language skills (vocabulary, syntax, narrative) that are the ultimate goal of intervention.

To illustrate another important point, recall that we summarized a relatively new, randomized controlled study that suggested that at least for police officers in D.C., the use of BWCs may not be helpful in reducing police violence and citizen complaints. The authors of the study were careful to make the point that there were other factors that may have influenced their results. Recall that all of the participants were from D.C, where the police receive high quality training and have participated in years of programs designed to reduce misconduct and improve the quality of their interactions with citizens. In this example, the use of BWCs with D.C. police may be comparable to the need for WM capacity training. The police in D.C. did not need the BWCs because the education and training programs they had in place were effective in bringing
about authentic positive outcomes for the skills the they needed to have. Similarly, when clinicians use good intervention strategies like those summarized in this paper (and other papers in this forum) there may not be a need for additional, specialized WM training in order for children to perform well in communicative and academic contexts. We recommend using authentic, functional strategies as the first option when selecting interventions that are likely to be maximally helpful for children with language disorders.

References


Nelwan, M. & Kroesbergen E. H. (2016). Limited near and far transfer effects of jungle memory working memory training on learning mathematics in children with attentional and


