LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-624993

Working Memory, Situation Models, and Synesthesia

G. A. Radvansky , B. S. Gibson, M. W. McNerney

March 7, 2013

Journal of Experimental Psychology: Learning, Memory, Cognition

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Working Memory, Situation Models, and Synesthesia

Gabriel A. Radvansky Bradley S. Gibson

University of Notre Dame

M. Windy McNerney

Lawrence Livermore National Laboratory

Please address all correspondence to:

G.A. Radvansky

Department of Psychology

218-C Haggar Hall

University of Notre Dame

Notre Dame, IN 46556

Phone: 574-631-6473

FAX: 574-631-8883

e-mail: Radvansky.1@nd.edu

Abstract

Research on language comprehension suggests a strong relationship between working memory span measures and language comprehension. However, there is also evidence to suggest that this relationship weakens at higher levels of comprehension, such as the situation model level. The current study explored this further by comparing ten grapheme-color synesthetes and 48 normal controls on a number of tests of complex working memory capacity and situation model level processing. On all tests of working memory capacity the synesthetes outperformed the controls. Importantly, there was no carry-over benefit for the synesthetes for processing at the situation model level. This reinforces the idea that while some aspects of language comprehension are related to working memory span scores, this applies less to situation model levels. This suggests that theories of working memory need to take into account this limitation and those working memory processes that are involved in situation model construction and processing need to be derived.

Key words: synesthesia, memory, comprehension

Working Memory, Situation Models, and Synesthesia

A great deal of research has been devoted to understanding how measures of working memory capacity relate to other aspects of human cognition. The many positive correlations that have been found and reported have even led some theorists to suggest that working memory span scores are strongly related to the construct of general intelligence (e.g., Conway, Kane, & Engle, 2003; Engle, Tuholski, Laughlin, & Conway, 1999). Thus, the meaning of these scores is of broad interest. Of particular focus here is the relationship between working memory span scores and language comprehension, with a particular emphasis on the situation model level. More specifically, we explored the degree to which larger working memory spans in synesthetes translates into benefits in processing at the situation model level.

Before going further, let's clarify the different levels of text representation, namely the surface form, textbase, and situation model levels (van Dijk & Kintsch, 1983). The surface form corresponds to the verbatim representation of a text to capture the exact words and/or syntax used. The textbase corresponds to the propositional idea units in a text apart from the specific wording. For example, a paraphrase is consistent with the textbase, even if it is not accurate at a verbatim level. Finally, the situation model (Johnson-Laird, 1983; Zwaan & Radvansky, 1998) represents what the text refers to, rather than the text itself (as is the case in surface form and textbase representations).

Working Memory Span and Language Comprehension

There have been many studies that have drawn a link between performance on complex working memory span tasks and performance on various language comprehension tasks. This line of work began with Daneman and Carpenter's (1980) seminal study that involved an exposition of the sentence span task. Essentially, in the sentence span task people are asked to read aloud a series of sentences, with increasing set sizes, and remember the last word from each one. Once the end of a sentence set is reached, a person is to recall these final words. Much of the work that has followed from this (e.g., Engle, Cantor, & Carullo, 1992; Friedman & Miyake, 2004) has largely followed the basic emphases of that study. In these sorts of studies, performance on a complex working memory span task, such as the sentence span or operation span measures is found to be correlated with performance on some index of language comprehension. The indices include such things as the verbal SAT, the ability to remember verbatim or propositional from a text, or the ability to engage in other text-based processing, such as identify anaphoric referents when explicitly probed for them.

Working Memory Span and Situation Model Processing

While working memory span scores can correlate with measures of language processing, often these measures of language processing involve processes at the surface form or textbase levels, such as memory for specific sentences in a text, or the meaning of particular vocabulary words. In some sense, finding a relationship between these measures is not all that surprising. Working memory span tasks often involve memory for single words and possibly the processing of single sentences. This is exactly the type of processing involved at the surface form and textbase levels.

There have been fewer studies looking at working memory span and language processing at the situation model level. There is some suggestion that there is not such a strong correspondence between working memory span and situation model level processing. For example, one study, by Radvansky and Copeland (2004b) failed to find a relationship between measures of working memory capacity, such as the reading span and operation span measures, and a number of situation model level processes, including sentence memory, causal structure, and inconsistency detection, among others.

Moreover, older adults, who typically have lower working memory span scores, often do not have processing difficulties at the situation model level (Radvansky & Dijkstra, 2007). Thus, working memory span scores may not be reliable indicators of the effectiveness of more complex cognitive processes, per se. That said, there is some evidence to suggest that there are some circumstances where declining working memory capacity with age can exert an influence. For example, Noh and Stine-Morrow (2009) found that older adults had more difficulty than younger adults in tracking multiple characters in a described event. Still, the more general finding is that the lower working memory span scores of older adults do not have a strong influence on situation model processing. Would the same hold true for people who are thought to be likely to score higher on working memory span tests, such as synesthetes? Would they show a carryover benefit to the situation model level or not?

Memory and Synesthesia

If a person experiences synesthesia, in addition to standard sensory experiences they have any number of inappropriate and involuntary sensory experiences (see

Grossenbacher & Lovelace, 2001; Hochel & Milán, 2008; Hubbard & Ramachandran, 2005; Rich & Mattingly, 2002 for reviews). This synesthesia is due to either a lower ability to suppress inappropriate feedback loops in perceptual processing (Grossenbacher & Lovelace, 2001) or an incomplete pruning of extra cortical connections during development (Maurer, 1997). Of particular concern here are those cases in which a person reports experiencing colors (photisms) when reading words, what is called grapheme-color synesthesia. Previous research with grapheme-color synesthetes has shown that memory for materials that elicit the synesthetic experience (letters and words) is superior to that of controls, either through subjective reports or through experimental verification (Luria, 1968; Mills, Innis, Westendorf, Owsianiecki, & McDonald, 2006; Smilek, Dixon, Cudahy, & Merikle, 2002; Ward, 2008; Yaro & Ward, 2007).

Recently, we (Gibson, Radvansky, Johnson, & McNerney, 2012; Radvansky, Gibson, & McNerney, 2011) have found that ten synesthetes had superior long-term memory for word lists compared to controls. Moreover, Gibson et al. reported that the synesthetes had superior simple memory span performance, although performance on complex working memory span measures (the ones that are correlated with measures of language processing) were not reported. These findings of overall superior memory for word lists and simple span tasks is consistent with the expectation that synesthetes would have higher complex working memory spans, than normal controls. More importantly, there was evidence that the synesthetes placed a greater emphasis on the words themselves compared to other factors. Specifically, the synesthetes had smaller or absent von Restorff effects when a list singleton was identified by a unique color or semantic meaning, and smaller DRM false memory effects. Thus, the synesthetes have superior memory for verbatim, item-specific information. However, it is not clear whether this benefit scales up and transfers to the situation model level.

Note that there is not a universal cognitive benefit of having synesthesia, but rather any performance enhancement is limited to those tasks that use materials that elicit the synesthetic experience. For example, as reported by Gibson et al. (2012), our synesthetes did not differ from normal controls on a spatial position task. Also, although it has not been formally reported, our synesthetes also did not differ from normal controls on the Shah and Miyake (1996) spatial span measure or the Shepard and Metzler (1971) mental rotation task. So, synesthesia provides no benefit to cognitive processes that do not involve the materials that elicit the synesthetic experience.

According to views that suggest that complex working memory span scores are indicators of general cognitive performance, grapheme-color synesthetes should have not only higher complex working memory span scores, but also superior verbal processing, even at the situation model level. Alternatively, there is also the possibility that the benefit synesthetes experience at surface and textbase levels will not scale up to the situation model level, consistent with other research showing that there is a weak relationship between working memory span scores and situation model level processes (Radvansky & Copeland, 2004b).

To distinguish between these possibilities, we analyzed data from synesthetes and controls on a number of complex working memory span tests, including the operation span, sentence span, and comprehension span measures. Next, we assessed the relationship between working memory span, and performance on a number of situation model processing tasks, including event indexing, levels of representation, and sentence memory. Finally, we explore whether there are processing benefits for synesthetes on a number of additional situation model level tasks, specifically target causal processing, temporal updating, and spatial integration.

Task Set 1: Working Memory Span Tests

The first series of tasks we address are complex working memory span measures. These tasks are of interest because our synesthetes experience colors when looking at letters and there is already some evidence to suggest that short-term memory for letters and words (Gibson et al., 2012), as well long-term memory for verbal materials, is better in synesthetes (Radvansky et al., 2011). As such, it is expected that they would perform better than normal controls on standard measures of verbal complex working memory.

Participants. Ten of the participants were synesthetes (eight female) who reported experiencing colors as they read different letters. These people were all students at Notre Dame, the same population from which the control subjects were drawn from. Three of synesthetes (one male and two female) also reported experiencing colors when listening to people's voices, and another reported experiencing taste sensations with some colors. Also, one of the synesthetes appears to have acquired her letter-color photisms in line with the color of the refrigerator magnets she had as a child, an origin that has also been reported elsewhere (Witthoft & Winawer, 2006). Another synesthete reported strong experiences of "male" and female" for the letters of the alphabet. This ordinal linguistic personification (OLP) has also been reported in other synesthetes (Simner & Holenstein, 2007; Simner & Hubbard, 2006; Smilek, Malcolmson, Carriere, Eller, Kwan, &

Reynolds, 2007). All of the synesthetes reported what would correspond to associative synesthesia rather than projector synesthesia (Dixon, Smilek, & Merikle, 2004). In associative synesthesia, the synesthetic sensory experience occurs as being in the mind of the synesthete, whereas in projective synesthesia, the sensory experience occurs as if it were out in the world.

All of the synesthetes were assessed using a computerized mapping task. In this task, people were presented with letters, digits and symbols on a computer screen. Each item was presented once in white on a black background and once in black on a white background in each of three blocks of trials. All of the items were randomly ordered within each block. The task was to select from a palette of 30 options the color that most closely corresponded to their synesthetic experience, if any. There was also a textbox provided to enter any comments the synesthetes may have had. This task obtained the experienced colors for each synesthete (needed for the colored word task).

This task also verified the synesthetic experience. We assessed consistency both across two testing sessions as well as within a given session. There were two measures of consistency. For the *strict criterion* the person needed to select the exact same option, but for the lenient criterion if a person selected two options within the same category (e.g., two shades of blue) they were scored as the same. We considered performance on the letters because we were primarily interested in verbal memory and comprehension here. Performance was consistent both within the sessions (strict: range = .74 to .97; *M* = .90: liberal: range = .90 to .99; *M* = .95), and across sessions (strict: range = .51 to .91; *M* = .72: liberal: range = .71 to .98; *M* = .88), which is similar to mean consistencies reported elsewhere (e.g., Hubbard & Ramachandran, 2005). As such, we are confident that our synesthetes' experiences were genuine and reasonably stable[1].

Forty-eight control subjects were drawn from the research participant pool in the Department of Psychology at the University of Notre Dame. As such, these control subjects were from the same population as the synesthetes. None of them reported having any synesthetic experiences.

Method. For the complex span tasks, three different tests were selected that included both retention and processing components. One complex span task was the Daneman and Carpenter (1987) reading span task. For this task people were presented with 2 to 6 sentences. When each sentence appeared on the screen the participant read it aloud. At the completion of the sentence, the experimenter pressed a button, which advanced the program to the next sentence. At the end of a set, the participant recalled the last word of each sentence in the order they were read. The experimenter typed the responses into the computer. People were aware that their memory for the final words would be tested.

A second complex span task was the Turner and Engle (1989) operation span task. For this task people were presented with 2 to 6 math problem-word pairs. Each math problem involved two steps, and was presented along with a solution. The task was to indicate whether the solution was correct. After the solution was evaluated, a word appeared. When each problem appeared the person read it aloud and gave an evaluation. At the completion of the problem, then the word was presented and read aloud. After this

[1] Many of the remaining mismatches for the color mapping process were adjacent colors, such as red and brown, or yellow and orange.

the experimenter pressed a button which advanced the program to the next item. When the end of a set was reached, the person recalled the words in the order they were read. The experimenter typed the responses into the computer.

Finally, a third complex span task was the Waters and Caplan (1996) comprehension span task. For this task people were presented with 2 to 6 sentences. Their task was to indicate whether each sentence was meaningful. An example of a meaningful sentence is “It was the website that the accountant accessed,” whereas a nonmeaningful sentence would be "The boss laughed at the raise that asked for an employee.” The task was to indicate whether the sentence was meaningful by pressing one of two buttons on the computer mouse. The left button was marked with a “Y” for “Yes, this is sensible,” whereas the right button was marked with an “N” for “No, this is not sensible.” After pressing the button, the computer advanced to the next sentence. When the end of a set was reached, the participant was to recall the last word of each sentence in the set in the order they were read. People typed their own responses into the computer, and only the current response was visible at one time.

Results and Discussion. Span tests were scored in two ways. One of these was the highest span level attained. For this scoring method, a person was given credit for achieving a span at a certain level if two or more sets were recalled at that level, and if one set was recalled then only half credit was given. For example, if a person recalled four sets at span level 4 before stopping, then the person’s span score would be 4. However, if a person recalled only one set at span level 4, then the person’s span score would be 3.5. The other scoring method was devised by Turner and Engle (1986). With this method, whenever a person correctly recalls a set, the person is awarded the same number of points as the size of that set. Then, all of the points are totaled. For example, if a person recalled three sets at level 2, and 1 set at level 3, then their span score would be 9 (i.e., 2 + 2 + 2 + 3). Statistics for these two measures are distinguished with the subscript *S* for the span level method, and *T* for the total score method.

The complex span test scores are reported in Table 1. There were significant effects of Group on the Sentence Span, $F_S(1,56) = 8.41$, $MSE = 1.1$, $p = .005$, $F_T(1,56) = 13.55$, $MSE = 218$, $p = .001$, Comprehension Span, $F_S(1,56) = 17.39$, $MSE = 1.3$, $p < .001$, $F_T(1,56) = 20.00$, $MSE = 180$, $p < .001$, and the Operation Span tests, $F_S(1,56) = 6.49$, $MSE = 1.5$, $p = .01$, $F_T(1,56) = 5.43$, $MSE = 528$, $p = .02$, with synesthetes substantially outperforming the control participants. Thus, it appears that the synesthetes were able to exploit their synesthetic experience to improve their ability to remember sets of words.

Task Set 2: Direct Comparison

Having established clear and large differences between synesthetes and normal controls on verbal complex working memory span tasks, we can now assess whether this processing boost carries over to the situation model level. At the time the working memory measures were gathered, three tasks were given that allow us to assess processing at the situation model level. One of these involved having people read complex texts, and then respond to recognition probes. The reading time data were submitted to Event Indexing regression analyses (Zwaan et al., 1995) that allowed for an assessment of the degree to which reading was influenced by text based characteristics, such as number of syllables and word frequency, as well as situation model characteristics, such as shifts in space or time.

The recognition probe data were submitted to a Schmalhofer and Glavanov (1987) analysis which uses signal detection measures to derive indices of the surface form, textbase, and situation model levels in memory. If synesthesia boosts memory at all levels of language comprehension, then synesthetes will outperform the normal controls on all three signal detection measures. However, if the benefits observed at the letter and word levels do not scale up to more complex levels of representation, then it is expected that the synesthetes will outperform the normal controls at the surface form and textbase levels, but not the situation model level.

Another task was a sentence memory task that used memory for single sentences in which distractor sentences could be altered on a surface form level, but not on the situation model level, or on both. For this task, people read a series of sentences, such as "Three turtles sat on a log and a fish swam beneath it". To assess the impact of situation models on making memory decisions, people were given a recognition test based on a procedure originally developed by Bransford and Franks (1971) and expanded on by subsequent research (Garnham, 1981; Jahn, 2004; Radvansky, Gerard, Zacks, & Hasher, 1990).

Of particular importance is whether the distractor items referred to the same situation as the original sentence. Half of these sentences had a ***confusable*** alternative version. For example, the original sentence "Three turtles sat on a log and a fish swam beneath it" can be readily confused with "Three turtles sat on a log and a fish swam beneath them" because they both describe the same situation. In contrast, the other half of the sentences had ***nonconfusable*** alternative versions. For example, the sentences "Three turtles sat next to a log and a fish swam beneath it" and "Three turtles sat next to a log and a fish swam beneath them" are less likely to be confused because they describe different events, even though the sentences have been changed in the same way as in the confusable version pairs. If people are using situation models to make their memory decisions, rather than a more verbatim or textbase representation, then error rates will be higher in the confusable than the nonconfusable condition. This is because in the confusable condition, both sentence versions map onto the same event, and so the same situation model would be used.

Method

Participants. The same participants as reported in Task 1 were tested in these tasks.

Materials and Procedure. For the Event Indexing analysis and Schmalhofer and Glavanov analyses, people read 12 stories. These stories were drawn from previous studies, including Radvansky, Zwaan, Curiel and Copeland (2001) Radvansky Copeland, & von Hippel (2010), and Narvaez, Radvansky, Lynchard & Copeland (2011). These texts were 54 to 85 (M = 70.5) sentences long and contained numerous event shifts of different types that can be assessed by the Event Indexing reading time analysis.

To assess memory at different levels of representation, people were given a recognition test based on a procedure developed by Schmalhofer and Galvanov (1986). This recognition test was composed of four types of probes; (1) ***Verbatim*** probes made of sentences that were actually in the text (e.g., "The plan was to dig a tunnel through the dividing wall."), (2) ***Paraphrase*** probes made of sentences that express the same propositional idea units that were in the text, but are worded differently (e.g., "The idea was to dig a passage through the dividing wall."), (3) ***Inference*** probes made of sentences that express ideas that people are likely to have inferred during reading, but which were never expressed directly (e.g., "The plotters wanted to dig a tunnel under the court house."), and (4) ***Incorrect*** probes made of sentences that are thematically consistent with the text, but are clearly incorrect (e.g., "The purpose of the tunnel was for transporting explosives.").

During reading, the texts were presented one sentence at a time on a computer, advancing to the next sentence by pressing the space bar. Reading times were recorded. After all of the stories were read, people were given a recognition test.

For the sentence memory task, people first read a series of 30 sentences derived from materials reported in other studies (i.e., Bransford & Franks, 1971; Garnham, 1981; Jahn, 2004; Radvansky, et al., 1990). As described earlier, half of these sentences had a ***confusable*** alternative version and half had ***nonconfusable*** alternative versions. These sentences were presented one at a time on a computer screen. People rated each sentence for pleasantness as an orienting task.

After reading all the original sentences, people were given a recognition test. For this test, all four sentence versions were presented in a pseudo-randomized list. The only constraint was that two versions of the same sentence could not follow one another. As each item was presented, people needed to indicate whether it was the sentence that had been read earlier, and they were warned that some of the altered sentences may differ only slightly from the original version. People responded by pressing one of two buttons on the computer mouse. They pressed the left button, marked with a “Y” for “Yes, I did read this sentence before,” and the right button marked with an “N” for “No, I did not read this sentence before.” Accuracy was the primary dependent measure in the confusable or nonconfusable conditions.

Results and Discussion

Event Indexing Analysis. For the reading time data, we applied the Event Indexing analysis (Zwaan et al., 1995). In this analysis, based on a procedure developed by Lorch and Myers (1990), the reading time data were submitted to a regression model that takes into account both text-based factors, including serial position, number of syllables, and word frequency, as well as situation model factors, including breaks in space, time, causality, intentionality, and entities.

The reading time data for each person were submitted to regression analyses with the text and situation model factors. The mean beta weights are reported in Table 2. For the text variables (number of syllables, serial position, and word frequency), the reading times of both the synesthetes and control participants were significantly affected. Moreover, there was a significant difference between the two groups in that the synesthetes were more affected by the number of syllables, $F(1,56) = 3.89$, $MSE = .011$, $p = .05$. However, there were no significant differences between these two groups on the other two variables, both ps > .50. For the situation model factors, the results were nearly identical for the control participants and synesthetes, and none of these differences were significant, all ps > .30. Thus, in general, although they have superior working memory span scores, there was no indication that the processing of event breaks in the described situation was affected by synesthesia. Thus, while synesthesia can influence processing at the surface form and textbase levels, consistent with performance on the verbal working memory span measures, it did not influence processing at the situation model level.

Schmalhofer and Glavanov analysis. For the recognition test, using the four probe types, signal detection analyses were done to gain strength estimates on three levels of representation. The simplest level is the ***surface form***, which is a mental representation of the actual information that was presented. A measure of surface form is gained by treating "yes" responses to Verbatim probes as hits and "yes" responses to Paraphrases as false alarms. These probes both convey information that was actually read, but differ in that only one was in the form that was encountered. The next level is the ***propositional textbase***, which is a mental representation of the idea units that were read apart from the form they took. A measure of this is gained by treating "yes" responses to Paraphrases as hits and "yes" responses to Inferences as false alarms. Neither of these probes was actually read, and they differ only in that one contains idea units that were actually read, whereas the other does not. Finally, there is the level of the ***situation model***, which serves as a mental simulation of a person's understanding of the functional relations among entities in the world. This includes both information that was read and inferences that were generated. A measure of situation models is gained by treating "yes" responses to Inferences as hits and "yes" responses to Incorrects as false alarms. Neither of these conveys ideas that were actually read, but they differ in that the Inferences contain information that is consistent with the described situation, whereas the Incorrects do not.

For the recognition test, the A' signal detection data are presented in Table 3. Separate between subjects ANOVAs were done for each level of representation. There was some indication of superior memory for the synesthetes at the surface form level, consistent with our other work, $F(1,56) = 3.56$, $MSE = .005$, $p = .06$. However, there were no differences at the textbase and situation model levels, both *F*s <1. Thus, while the synesthetes do have marginally better verbatim memory, they have no advantage or deficit at the more abstract levels of memory, including the textbase level. Like the reading time data, this is consistent with the idea that synesthesia working memory span benefits do not carry over to more complex levels of thought.

Sentence Memory task. The accuracy data for the sentence versions were submitted to a 2 (Group) X 2 (Condition: confusable vs. nonconfusable) mixed ANOVA. The main effect of Group was significant, $F(1,56) = 3.97$, $MSE = .041$, $p = .05$, with synesthetes being more accurate ($M = .54$) than the controls ($M = .44$). This is consistent with the general finding that synesthetes have superior overall surface form, verbatim memory compared to the controls.

There was also a main effect of Condition, $F(1,56) = 35.41$, $MSE = .015$, $p < .001$, with people being less accurate in the confusable condition ($M = .40$) than the nonconfusable condition ($M = .58$). Importantly, the interaction was not significant, $F < 1$, with a similar difference between the conditions for the synesthetes (confusable = .44; nonconfusable = .64) and the controls (confusable = .36; nonconfusable = .52). Thus, both groups performed similarly in terms of the nature of the effect of confusability. The verbatim memory benefit synesthetes experience for individual words did not influence sentence memory in terms of the degree of involvement of situation models in making memory decisions. As such, this pattern of data is most consistent with the idea that synesthesia may benefit lower level processing, but that this benefit does not necessarily scale up to the situation model level.

Relations to working memory span scores. Although the synesthetes have higher working memory span scores, and there was no difference between synesthetes and controls on the situation model measures, to further explore the relation between working memory capacity and situation model processing we calculated correlations between each of our complex span measures and the various task measures.

For the reading time data, for the sentence span task, there were significant correlations with number of syllables, $r = -.29$, $p = .03$, and word frequency, $r = .30$, $p = .02$. For the comprehension span task, there was a significant correlation with word frequency, $r = .27$, $p = .04$. For the operation span, there were no significant correlations. Thus, as found in previous research, there is a correspondence between working memory span scores and some aspects of language comprehension. However, this was confined to the influence of the number of syllables and word frequency. There were no significant correlations with serial position (all ps $\geq .15$), which some researchers have suggested reflects some situation model level processes (e.g., Stine-Morrow, Loveless, & Soederberg, 1996). More importantly, there were no significant correlations with any of the situation model factors (all ps $\geq .17$), consistent with the idea that working memory span scores do not strongly capture the mental processes operating at the situation model level.

For the recognition test data, there were significant correlations of comprehension span, $r = .32$, $p = .02$, and operation span scores, $r = .26$, $p = .05$, and a marginally significant correlation with the sentence span score, $r = .24$, $p = .07$. This is consistent with the idea that working memory span tests tap into word level processing, which is not surprising because they often involve people remembering individual words outside of context. There were no significant correlations of any of the working memory span measures and scores at the textbase level, (all ps $\leq$ 12), or the situation model level (all ps $< .10$). This pattern of data parallels that reported by Radvansky and Copeland (2004b).

Finally, looking at the sentence memory task, for the comprehension span scores, there were marginally significant correlations with rate of rejecting confusable foils when originally given the nonconfusable sentence version, $r = .23$, $p = .08$, or vice versa, $r = .25$, $p = .06$. There were no significant correlations for the sentence span or operation span scores (all ps $> .12$). Again, the relatively sparse correlation with performance on this measure is consistent with the idea that working memory span scores are not tapping strongly into more complex language comprehension processes.

Task Set 3: Indirect Comparison

The previous task set involved a number of tasks that assessed processing at the situation model and other levels, and were collected at the same time as the working memory span measures. In addition to these, at another period of time, we collected performance measures on a number of less general, more targeted, situation model measures. The same synesthetes were tested, but, due to practical constraints (it was a different school year), different normal control participants were involved. So, while we do not specifically have working memory span measures for those participants, as was made clear in Task Set 1 the synesthetes have much larger working memory spans than would be expected in the normal population. So, the assumption that our synesthetes have larger working memory spans compared to this other group of normal controls is a near certainty. Given this, is there any evidence that the synesthetes could outperform the normal controls on these other situation model level measures?

As people read and comprehend, the described events are often changing. For comprehension to be successful, people need to successfully update and track information in their situation models. This updating can occur along any of a number of event dimensions (Zwaan et al., 1995). As such, we assessed whether synesthesia affects the updating along some of these. The focus of the measures in Task Set 3 is on situation model processing for causal, temporal, and spatial event dimensions.

In terms of the causal dimension, when people encounter a causal break in a text, such as being told an effect with the cause, people then need to infer the missing causal information. Thus, this is an event boundary, and it takes people some time to update their situation model as reflected by reading times (Singer, Halldorsorn, Lear, & Andrusiak, 1992). If the verbal working memory benefit associated with synesthesia carries over to processing at the situation model level, then it is expected that synesthetes would be more likely to draw such causal connections as they would have more information available in working memory, allowing them to notice and draw the causal links. However, if increased verbal working memory spans do not capture this type of mental process, then there will be no difference between the synesthetes and controls.

Related to the idea that causal connections are important for processing, we can also look at the influence of causal relations on other types of information. Here we focus on spatial relations. Spatial relations that play some causal role are called functional (Radvansky & Copeland, 2001). Because of this additional causal element, when functional spatial relations are encountered in a text, they are read more quickly than nonfunctional spatial relations. Again, if verbal working memory span scores reflect an overall increase in language processing, then it is expected that synesthetes will outperform the controls, but if it does not then it is expected that there will be no difference.

Other aspects of events that people need to track while reading are changes in time, such as a change from one time frame to another. To assess temporal updating we used a procedure developed by Zwaan (1996) in which, during the course of reading, a critical event occurred and then there was a temporal boundary marked by either the phrase “a moment later” which was essentially no boundary, and “a day later” which is a substantial boundary in the context of these texts. After this, people were probed for the action the main character was doing prior to the event shift. These actions were all temporally limited and would have been completed long before the event boundary, such as drinking a cup of coffee. Again, if the increased verbal working memory capacity spills over into higher levels of processing, then it is expected that synesthetes would be more sensitive to such changes as they would be more likely to notice and process this information along with all of the other information in the story. However, if there is no additional benefit outside of the individual word level, then it is expected that there will be no difference between the synesthetes and the controls.

To assess spatial processing at the situation model level we looked at the integration of spatial layout information into a situation model. This task was patterned after a study by Mani and Johnson-Laird (1982), in which people were given with a series of descriptions of the spatial arrangement of four objects, presented one sentence at a time in either a continuous or discontinuous form. An example of one description in both forms is presented below.

Continuous

The hammer is in front of the screwdriver.

The screwdriver is to the left of the saw.

The saw is in front of the wrench.

Discontinuous

The hammer is in front of the screwdriver.

The saw is in front of the wrench.

The screwdriver is to the left of the saw.

After reading, people indicated which of eight test diagrams corresponded to that description. Performance is better in the continuous condition because the second and third sentences always refer to entities that were mentioned before. However, in the discontinuous version, people need to hold the first two sentences in mind and cannot integrate them until the third sentence. Performance is scored as the number of diagrams correctly identified.

Moreover, in terms of the sentence reading times, the typical pattern is for reading times to decrease, particularly from sentences 2 to 3, as people move through the descriptions, consistent with the idea that people are simply adding new information to an existing situation model. In comparison, for the discontinuous condition, there is an increase in reading times from sentences 2 to 3. This is because this is the point at which people can integrate the information from the prior two sentences, as well as the new information in the third sentence. If verbal working memory scores reflect broad-based increased processing effectiveness, then it is expected that synesthetes will show a smaller difference between the continuous and discontinuous condition. This is because they would be better able to maintain their memory for the two sentences more precisely in the discontinuous condition to allow them to then integrate all of the information properly when the third sentence is encountered. However, if the increased working memory span found in our synesthetes does not extend beyond the individual word level. Then no difference between the synesthetes and controls will be found.

Method

Participants. The same synesthetes were used. However, because data on these tasks were done after we obtained the other results, a different set of 48 control participants were used from the same population. Again, these control participants were from the same population as the synesthetes. None of them reported having any synesthetic experiences

Materials and procedure. For this task set, there were four tasks. Two of these addressed the processing of causal information. These were Singer et al.'s (1992) causal connection task, and Radvansky and Copeland's (2001) functional relations task. Finally, another task was aimed a tracking temporal updating and was modeled after work by Zwaan (1996).

Causal Connections. For the causal updating task we used a version of a task developed by Singer et al. (1992). In this task people read 64 two sentence descriptions. What was manipulated was whether there was a causal relation between the two. An example of a causally related pair is:

Dorothy poured the bucket of water on the bonfire.

The fire went out.

In this case, it is clear that the second sentence is causally related to the first. In this case, both the causal antecedent and consequent are directly provided. In comparison, the following is an example of a causally unrelated pair:

Dorothy placed the bucket of water by the bonfire.

The fire went out.

Here, people need to infer the causal antecedent, which is that the water was poured in the fire. Thus, reading time for the second sentence should take longer in this case than when the two are directly causally related, and it is easier for people to integrate the information into a situation model. The absence of a direct causal relation in the text creates a causal break. For a given person, 12 of the passages conveyed a causal relation, 12 were non-causal, and 40 were filler passages.

After each two-sentence passage, people are given a probe question to assess the activation of causal information from semantic memory. For the examples above, the probe would be “Does water extinguish fire?”

Functional Relations. For functional spatial relation task, we used a paradigm developed by Radvansky and Copeland (2001) in which people read 16 passages that contained descriptions of spatial relations, along with two practice texts. For half the texts, these spatial relations were functional. That is, the particular spatial relation allowed two objects to be related in some way. For example, if a person were standing under a bridge, this would keep them dry from the rain. On the other half of the trials, the spatial relations were nonfunctional. That is, although two entities were in some spatial relation to one another, it was unlikely that they were interacting. For example, a person standing under a streetlamp would still get wet in the rain. We recorded reading times for the spatial sentences. After reading, people were given a four alternative forced-choice recognition test. The spatial relation that was read earlier was presented along with three distracter sentences. The task was to select the sentence that was read earlier. Response times and accuracy on these trials was also recorded.

Temporal Shifts. To assess temporal updating, people read a series of 36 texts, based on a study by Zwaan (1996). Half of these texts contained critical temporal shift sentences, and half were fillers. Embedded in the critical texts were sentences that described temporal shifts. On half of the critical trials it was a relatively short period of time, such as “a moment later” in which time is more continuous, whereas for the other half there was a long temporal shift such as “a day later.” Reading times were recorded.

Also, immediately after the critical sentences, a probe was presented that described the activity of the story protagonist just prior to the temporal shift. If there is a short time shift, then it is highly likely that the character would still be involved in that activity. However, after a long time shift, it is likely that the activity would have stopped. For example, if a person reads that “Chris started opening the mail. A moment later, Chris felt a headache.” it is highly likely that Chris is still opening the mail. However, if the second sentence were “A day later, Chris felt a headache.” it is unlikely that Chris will still be opening the mail. As such, responses to the probe should be faster and more accurate in the short temporal condition than the long temporal condition.

Spatial Integration. People were given 20 descriptions of the spatial arrangement of four objects, following Mani and Johnson-Laird (1982). These descriptions were presented on a computer, once sentence at a time. After reading each description, people were shown eight possible configurations of the objects and asked to select which figure corresponded to the description. This was done by selecting a radio button that corresponded to a particular diagram on the computer screen. Reading times were collected along with response times and accuracy rates for the selection of the spatial arrangement alternatives.

Results and Discussion

Causal Connections. The reading time data for the critical second sentence, summarized in Table 4, were submitted to a 2 (Group) X 2 (Condition: causal vs. related) mixed ANOVA. There was a main effect of Group, $F(1,56) = 6.49$, $MSE = 149093$, $p = .01$, with synesthetes reading the sentences more slowly than the controls. There was also a main effect of Condition, $F(1,56) = 9.20$, $MSE = 19189$, $p = .004$, with people reading faster for the causal condition than the related condition. This is the causal updating effect. The interaction was not significant, $F < 1$, suggesting no hint of a difference between the synesthetes and the controls in the ability to update causal information, despite the synesthetes better than normal working memory span scores.

For the casual probe task, the response time and accuracy data (shown in Table 5) were submitted to 2 (Group) X 2 (Condition) mixed ANOVAs. For the response times, the main effect of Group was not significant, $F(1,56) = 2.35$, $MSE = 670470$, $p = .13$, but the main effect Condition was marginally significant, $F(1,56) = 2.88$, $MSE = 158682$, $p = .09$, demonstrating the basic causal updating effect. The interaction was not significant, $F < 1$, again suggesting no difference between the synesthetes and controls in the updating of causal information. None of the effects were significant in the accuracy analysis, all *F*s < 1. Overall, again, there is no additional benefit afforded to the synesthetes for the processing of this type of information.

Functional Relations. The reading time data, summarized in Table 4, were submitted to a 2 (Group) X 2 (Condition: functional vs. neutral) mixed ANOVA. There was a main effect of Group, $F(1,56) = 7.35$, $MSE = 797211$, $p = .009$, with synesthetes reading more slowly than the controls. There was also a main effect of Condition, $F(1,56) = 20.52$, $MSE = 152486$, $p < .001$, with people reading the functional versions faster than the nonfunctional versions. This is the functionality effect. The interaction was not significant, $F(1,56) = 1.43$, $MSE = 152486$, $p = .23$, suggesting that synesthetes and controls are similarly sensitive to this type of information despite the synesthetes better than normal working memory span scores.

The recognition data are summarized in Table 5. For the recognition response time data, there were no significant effects of Group, $F(1,56) = 1.22$, $MSE = 821100$, $p = .27$, Condition, $F = 1.00$, or the interaction, $F(1,56) = 2.17$, $MSE = 333188$, $p = .15$. Similarly, for the accuracy data, neither main effects were significant, both *F*s < 1, nor was the interaction, $F(1,56) = 1.26$, $MSE = .019$, $p = .27$. So, overall there were no benefits or costs of having synesthesia on these memory tasks.

Temporal Shifts. The reading time, probe response time and error rate data were submitted to 2 (Group) X 2 (Condition: short vs. long) mixed ANOVAs. For the reading time data, summarized in Table 3, the main effect of Group was not significant, $F(1,56) = 2.21$, $MSE = 265068$, $p = .14$. However, the main effect of Condition was, $F(1,56) = 5.64$, $MSE = 85297$, $p = .02$, with people reading the critical sentences slower when there was a long temporal shift ($M = 1482$ ms per sentence) than a short one ($M = 1355$ ms per sentence). This is the temporal updating effect. The interaction was not significant, $F < 1$, suggesting that both groups updated temporal information in a similar manner, again despite the synesthetes better than normal working memory span scores.

For the probe response time data (shown in Table 5), neither the main effect of Group, $F < 1$, Condition, $F(1,56) = 1.02$, $MSE = 205542$, $p = .32$, nor the interaction were significant, $F(1,56) = 2.06$, $MSE = 205542$, $p = .16$. Similarly, for the accuracy data, neither the main effect of Group, $F < 1$, Condition, $F(1,56) = 1.05$, $MSE = .006$, $p = .31$, nor the interaction were significant, $F(1,56) = 1.61$, $MSE = .006$, $p = .21$.

Spatial Integration. The reading time data, shown in Table 6, were submitted to a 2 (Group) X 2 (Condition: continuous vs. discontinuous) X 3 (Sentence number) mixed ANOVA. There was a significant effect of Group, $F(1,56) = 5.57$, $MSE = 22007363$, $p = .02$, with the synesthetes reading slower ($M = 6554$ ms / sentence) than the controls ($M = 4982$ ms / sentence). Although there was no main effect of Condition, $F < 1$, there was a significant effect of Sentence, $F(2,112) = 4.31$, $MSE = 2820193$, $p = .02$, which was modified by a Condition X Sentence interaction, $F(2,112) = 11.25$, $MSE = 2007443$, $p < .001$. The data from the continuous condition showed a decrease in reading time from Sentences 2 to 3 as the information could be easily integrated into the situation model, whereas the for discontinuous condition there was an increase in reading time at the third sentence, which is where integration across the three sentence types would have occurred.

In addition, there was a marginally significant interaction of Group and Condition, $F(1,56) = 3.62$, $MSE = 2417853$, $p = .06$. The reading time data suggest that the synesthetes were not processing the information similar to the controls, especially in the discontinuous condition, which does not show as much evidence of reading time changes as a function of the amount of situation model integration that is needed. For the control subjects, the main effect of Condition was significant, $F(1,47) = 7.49$, $MSE = 1959080$, $p = .009$. This effect was not significant for the synesthetes, $F < 1$. So, this suggests that the synesthetes are not as sensitive to the developing event-level representation of the situation model on this task, perhaps because they may be more reliant on their verbatim and textbase representations.

The recognition test response time and accuracy data are shown in Table 5. These data were submitted to 2 (Group) X 2 (Condition: continuous vs. discontinuous) mixed ANOVAs. Five control participants were excluded from the response time analyses because they had no correct responses for either or both of the conditions. For the response time data, the synesthetes were slower than the controls, $F(1,51) = 5.27$, $MSE = 15706007$, $p = .03$. Neither the main effect of Condition, $F(1,51) = 2.27$, $MSE = 11434780$, $p = .14$, nor the interaction were significant, $F < 1$.

The lack of a significant main effect of Condition is surprising. To explore this, although the interaction was not significant, we reanalyzed the data for the control subjects alone. This analysis did reveal a marginally significant effect of Condition, $F(1,42) = 3.65$, $MSE = 9906010$, $p = .06$, with people responding more slowly in the

Discontinuous condition that the Continuous condition. Thus, ease of integration does appear to be having some impact on later memory. The lack of a main effect of Condition in the complete analysis appears to be a high degree of variability in the data from the synesthetes. When the synesthetes were considered separately, the effect of condition was not significant, $F < 1$. Overall, this suggests that this is an aspect of comprehension that is worthy of more in-depth study. Note that these interpretations should be taken with caution as we cannot draw any firm conclusions at this point.

For the accuracy data, the synesthetes were more accurate than the controls, $F(1,56) = 4.34$, $MSE = .162$, $p = .04$. While the pattern of accuracy data suggests that there was a difference between the Continuous and Discontinuous conditions, neither the main effect of Condition, $F(1,56) = 1.07$, $MSE = .024$, $p = .31$, nor the interaction were significant, $F < 1$. As with the response time data, the absence of significant main effect of Condition again led us to wonder whether our manipulation was strong enough. When the data from the synesthetes was excluded, the control participants revealed a main effect of condition, $F(1,42) = 3.92$, $MSE = .026$, $p = .05$, with people being less accurate in the Discontinuous condition than the Continuous condition. When the data from the synesthetes are considered alone, the effect was not significant, $F < 1$. Again, as with the response time data, the integration of information to update and form a situation model would be worthy of further study.

So, at this point, like Task Series 2, there is no evidence to suggest that there is a strong benefit for the synesthetes and our normal controls in the processing of information at the situation model level. This is despite the fact that the synesthetes have been shown to have much larger than normal complex working memory span scores. This is further support for the idea that while grapheme-color synesthesia can boost performance on some memory measures, such as our working memory span tests, there is not an overall verbal memory benefit. Moreover, these findings are also consistent with the idea that higher performance on working memory span tests does not necessarily capture performance abilities at the situation model level of processing.

General Discussion

The results of the study reported here further demonstrate that while working memory span scores do predict important aspects of human cognition, there are also some limitations. One of these, consistent with other research (Radvansky & Copeland, 2004) is the ability of such scores to capture more complex types of thinking, such as language processing at the situation model level. In this study, we found that people with grapheme-color synesthesia have higher complex working memory span scores. Consistent with this benefit, our synesthetes outperformed our normal controls on measures that focus on lower levels of language processing, such as the influence of word frequency on reading and memory for text at the surface form level.

However, despite this large advantage, there was no evidence that this benefit carried over to more complex types of language processing. In most of the tasks used here, while there was evidence of an influence of synesthesia on processing at the surface form level, there was no influence on the situation model level. The one exception to this general pattern was the spatial integration task, which suggested that synesthesia may have a small, but *negative* effect, on the amount of processing done at the situation model level.

An unexpected result of these studies was the near universal slower overall reading times for the synesthetes, across a range of tasks. In normal reading, people draw on the abstract textbase and situation model levels to facilitate comprehension. However, if a person were focused more on the word level, as appears to be the case with our synesthetes, then these other levels may exert a smaller influence, and would not facilitate the speed of processing during reading to as great a degree as with normal controls. That is, our normal controls may have a greater dependency on their situation models during reading that would actually speed their processing. In comparison, the synesthetes may have been slowed because of more intensive processing at lower level of processing, such as with the specifics of individual words. This is an interesting avenue for further research.

So, overall, while there are substantial benefits of synesthesia to memory, especially in terms of standard working memory measures for verbal materials, there is no overall boost in performance beyond the level of the individual words. In some sense this parallels the work on aging and language comprehension and memory in that older adults typically score lower on working memory span measures, but do not show processing deficits at the situation model level (Radvansky & Dijkstra, 2007). Although there has been some research suggesting that working memory capacity is related to higher-level cognitive functioning (e.g., Engle, Cantor, & Carullo, 1992; Engle, Tuholski, Laughlin, & Conway, 1999; Unsworth & Engle, 2007), there is also some evidence to the contrary (e.g., Radvansky & Copeland, 2001; 2004a; 2004b), including evidence that the reduced working memory capacities of older adults do not translate in similar deficiencies at the situation model level (Radvansky & Dijkstra, 2007).

The memory and processing benefits from synesthesia do not scale up to more complex forms of processing, and traditional working memory span scores are not reliable indicators of event level processes, at least not directly. It is certainly the case that more complex levels of language processing depend on successful processing at lower levels. As such, if there are serious problems at the surface form level of language processing, which working memory span scores seem to reflect, then processing at the situation model level is going to be hurt For example, if a story provides a description of two characters, Mary and Elizabeth, it will be necessary to maintain the names of these characters in working memory, to help keep track of who is who. If working memory span is reduced, then performance will suffer at the situation model level as the comprehender will have difficulty properly assigning new information to one character or the other.

Conclusions

Across a range of memory and comprehension tasks we compared the performance of ten synesthetes with normal controls. The data were largely consistent with the idea that working memory span scores do not predict higher-level processes well. The influence of grapheme-color synesthesia on cognition was evidence as superior performance for synesthetes for those materials that elicited the synesthetic experience, and that could be improved by word level of processing. There was no evidence of better processing on the part of the synesthetes at the situation model level. Overall, either no difference was observed between these two groups, or the synesthetes had effects that actually showed some processing costs. Thus, for grapheme-color synesthetes, such as those studied here, there are clear memory benefits at the word level, but this benefit does not scale up and influence processing at the situation model level. So, while it is clear that working memory span scores do tell us some important things about language processing and memory, there are aspects of comprehension and memory, such as the situation model level, that is not well indexed by such measures.

Acknowledgements:

Prepared at LLNL under Contract DE-AC52-07NA27344.

References

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331-350.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7,* 547-552.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19,* 450-466.

Dixon, M. J., Smilek, D., & Merikle, P. M (2004). Not all synesthetes are created equal: Projector versus associative synesthetes. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 335-343.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 972-992.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309 –331.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory & Language, 51,* 136-158.

Garnham, A. (1981). Mental models as representations of text. *Memory & Cognition, 9,* 560-565.

Gibson, B. S., Radvansky, G. A., Johnson, A., & McNerney, M. W. (2012). Grapheme-color synesthesia can enhance immediate memory without disrupting the encoding of relational cues. *Psychonomic Bulletin & Review, 19,* 172-177.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Science, 5,* 36-41.

Hochel, M., & Milán, E. G. (2008). Synaesthesia: The existing state of affairs. *Cognitive Neuropsychology, 25,* 93-117.

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron, 48,* 509-520.

Jahn, G. (2004). Three turtles in danger: Spontaneous construction of causally relevant spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 969-987.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Lorch, R. F., & Myers, J. L. (1990). Regression analysis of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149– 157.

Luria, A. (1968). *The Mind of a Mnemonist*. New York: Basic Books.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10,* 181-187.

Maurer, D. (1997). Neonatal synaesthesia: implications for the processing of speech and faces. In *Synaesthesia: Classic and Contemporary Readings*, S. Baron-Cohen, & J.E. Harrison (Eds.), pp. 224–242 Blackwell: Malden, MA.

Mills, C. B., Innis, J., Westendorf, T., Owsianiecki, L., & McDonald, A. (2006). Effect of a synesthete's photisms on name recall. *Cortex, 42*, 155-163.

Narvaez, D., Radvansky, G. A., Lynchard, N. A., & Copeland, D. E. (2011). Are older adults more attuned to morally-charged information? *Experimental Aging Research, 37,* 398–434.

Noh, S. R., & Stine-Morrow, E. A. L. (2009). Age differences in tracking characters during narrative comprehension. *Memory & Cognition*, *37*, 769– 778.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17,* 273-281.

Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition, 29,* 1073-1080.

Radvansky, G. A., & Copeland, D. E. (2004a). Reasoning, integration, inference alteration and text comprehension. *Canadian Journal of Experimental Psychology, 58,* 133-141.

Radvansky, G. A., & Copeland, D. E. (2004b). Working memory and situation model processing: Language comprehension and memory. *American Journal of Psychology, 117*, 191-213.

Radvansky, G. A., Copeland, D. E. & von Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology, 46,* 51-60.

Radvansky, G. A. & Dijkstra, K. (2007) Aging and situation model processing. *Psychonomic Bulletin & Review, 14,* 1027-1042.

Radvansky, G. A., Gerard, L. D., Zacks, R. T., & Hasher, L. (1990). Younger and older adults' use of mental models as representations for text materials. *Psychology and Aging, 5,* 209-214.

Radvansky, G. A., Gibson, B. S., & McNerney, M. W. (2011). Synesthesia and memory: Color congruency, von Restorff and false memory effects. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 37,* 219-229.

Radvansky, G. A., Zwaan, R. A., Curiel, J. M., & Copeland, D. E. (2001). Situation models and aging. *Psychology and Aging, 16,* 145-160.

Rich, A. N., & Mattingly, J. B. (2002). Anomolous perception in synaesthesia: A cognitive neuroscience perspective. *Nature Reviews: Neuroscience, 3,* 43-51.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory & Language*, *25,* 279- 294.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125,* 4-27.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701– 703.

Simner, J., & Holenstein, E. (2007). Ordinal linguistic personification as a variant of synesthesia. *Journal of Cognitive Neuroscience, 19,* 694–703.

Simner, J., & Hubbard, E. M. (2006). Variants of synesthesia interact in cognitive tasks: Evidence for implicit associations and late connectivity in cross-talk theories. *Neuroscience, 143,* 805–814.

Singer, M., Halldorsorn, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31,* 507-524.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P.M. (2002). Synesthetic color experiences influence memory. *Psychological Science, 13,* 548-552.

Smilek, D., Malcolmson, K. A., Carriere, J. S., Eller, M., Kwan, D., & Reynolds, M. (2007). When "3" is a jerk and "E" is a king: Personifying inanimate objects in synesthesia. *Journal of Cognitive Neuroscience, 19,* 981–992.

Stine-Morrow, E. A. L., Loveless, M. K., & Soederberg, L. M. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology & Aging*, *11,* 475-486.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language, 28,* 127-154.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038-1066.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies in Discourse Comprehension*. New York: Academic Press.

Ward, J. (2008). *The Frog Who Croaked Blue*. New York: Routledge.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology, 49A,* 51-74.

Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex, 42,* 175–183.

Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: What is superior about the memory of synaesthetes? *Quarterly Journal of Experimental Psychology, 60,* 681-695.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22,* 1196-1207.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21,* 386-397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123,* 162-185.

Table 1. Mean performance on the complex span tests for both highest span level, and total correct scorings (standard errors in parentheses).

Highest Span Level

	Synesthetes	Controls
Sentence Span *	4.3 (0.4)	3.2 (0.1)
Comprehension Span *	5.6 (0.4)	4.0 (0.2)
Operation Span *	4.6 (0.4)	3.5 (0.2)

Total Correct

	Synesthetes	Controls
Sentence Span *	42 (7.8)	23 (1.7)
Comprehension Span *	44 (4.9)	23 (1.9)
Operation Span *	49 (7.7)	30 (3.3)

Table 2. Performance on the event indexing regression analysis in standardized beta weights.

	Syll.	Serial	Freq.	Space	Time	Causal	Goals	Entity
Synesthetes	.58 *	-.07 *	-.05 *	.02 *	.00	.05 *	.03 *	.04 *
Control	.58 *	-.07 *	-.05 *	.02 *	.00	.05 *	.03 *	.02 *

* significantly different from 0, $p < .05$

Table 3. A' scores for the different levels of representation for memory for information presented in stories (standard errors in parentheses).

	Surface Form	Textbase	Situation Model
Synesthetes	.70 (.02)	.76 (.02)	.71 (.02)
Controls	.65 (.01)	.75 (.01)	.70 (.01)

Table 4. Mean reading time performance on the event updating tasks (standard errors in parentheses). For the various tasks, the following were the condition assignments for the Experimental condition: Causal = causal; Functional = functional; Time = long; space = dissociated.

	Experimental	Control
Causal		
Synesthetes	1039 (103)	1186 (143)
Controls	807 (34)	886 (46)
Functional		
Synesthetes	1983 (278)	2533 (466)
Controls	1503 (66)	1823 (87)
Time		
Synesthetes	1692 (210)	1456 (165)
Controls	1438 (63)	1334 (44)

Table 5. Performance on the event updating tasks (standard errors in parentheses). For the various tasks, the following were the condition assignments for the Experimental condition: Causal = causal; Functional = functional; Time = long; Space = dissociated; Integration = continuous.

	Response times		Accuracy	
	Experimental	Control	Experimental	Control
Causal				
Synesthetes	2011 (237)	2198 (320)	.94 (.02)	.94 (.02)
Controls	1724 (75)	1869 (93)	.95 (.01)	.96 (.01)
Functional				
Synesthetes	2582 (239)	2933 (234)	.82 (.09)	.86 (.02)
Controls	2545 (116)	2477 (103)	.82 (.02)	.80 (.02)
Time				
Synesthetes	1462 (79)	1734 (165)	.93 (.03)	.90 (.02)
Controls	1549 (87)	1502 (80)	.90 (.01)	.91 (.01)
Integration				
Synesthetes	9519 (1595)	10753 (781)	.79 (.09)	.77 (.08)
Controls	7230 (426)	8527 (654)	.61 (.05)	.54 (.04)

Table 6. Reading time data for the integration task (standard errors in parentheses).

	Continuous			Discontinuous		
	1	2	3	1	2	3
Synesthetes	6762 (792)	7100 (579)	5218 (547)	6897 (799)	6381 (568)	6968 (934)
Controls	5670 (318)	5748 (365)	4207 (312)	5267 (326)	4178 (291)	4825 (446)