Synesthesia and Memory: Color Congruency, von Restorff, and False Memory Effects

Gabriel A. Radvansky, Bradley S. Gibson, and M. Windy McNerney
University of Notre Dame

In the current study, we explored the influence of synesthesia on memory for word lists. We tested 10 grapheme-color synesthetes who reported an experience of color when reading letters or words. We replicated a previous finding that memory is compromised when synesthetic color is incongruent with perceptual color. Beyond this, we found that, although their memory for word lists was superior overall, synesthetes did not exhibit typical color- or semantic-defined von Restorff isolation effects (von Restorff, 1933) compared with control participants. Moreover, our synesthetes exhibited a reduced Deese–Roediger–McDermott false memory effect (Deese, 1959; Roediger & McDermott, 1995). Taken as a whole, these findings are consistent with the idea that color-grapheme synesthesia can lead people to place a greater emphasis on item-specific processing and surface form characteristics of words in a list (e.g., the letters that make them up) relative to relational processing and more meaning-based processes.

Keywords: synesthesia, memory, isolation effect, false memory

Synesthesia and Memory

Synesthesia is a condition in which a person reports inappropriate and involuntary sensory experiences in addition to the standard ones (for reviews, see Grossenbacher & Lovelace, 2001; Hochel & Milán, 2008; Hubbard & Ramachandran, 2005; Rich & Mattingly, 2002). For example, a person may report experiencing colors (photisms) when reading words, a condition called grapheme-color synesthesia. These letter–color associations are unique to individuals, although there are regularities across large numbers of synesthetes (Simner et al., 2005). This is also one of the more common types of synesthesia, occurring in 45%–65% of synesthetes (Day, 2009; Simner et al., 2006). The focus of the current work is on how synesthesia may influence memory performance for lists of words that elicit the synesthetic experience.

Causes, Experience, and Influence of Synesthesia

A number of studies have looked at the experience and neurological causes of synesthesia. One idea is that synesthetes do not have a set of unique neural connections but lack the ability to sufficiently suppress inappropriate feedback loops in perceptual processing (Grossenbacher & Lovelace, 2001). Another is that the brain of a synesthete is functionally distinct because of an incomplete pruning of extra cortical connections during development (Maurer, 1997).

Regardless of the underlying cause, it is still clear that synesthetes demonstrate a biological uniqueness in information processing. Some functional magnetic resonance imaging studies have revealed that synesthetes show greater activation in the left medial lingual gyrus during their synesthetic experiences (Rich et al., 2006) as well as unique interactions between areas V4 and/or V8 with the fusiform gyrus during perceptual tasks (Hubbard, Arman, Ramachandran, & Boynton, 2005; Ramachandran & Hubbard, 2001). Given that these findings were restricted to regions of the brain associated with higher levels of visual processing, and not in the lower visual processing regions (such as V2, although Aleman, Rutten, Sitkoorn, Dautzenberg, & Ramsey, 2001, reported activation in area V1 for one synesthete), this suggests that the synesthetic experience is unique at the level of perception rather than initial sensation.

Further evidence for this has been found with event-related potential (ERP) recordings. An enhanced positivity over the frontal region of the scalp, beginning at 200 ms after stimulus onset and lasting for several hundred ms, has been reported with synesthesia (Grossenbacher & Lovelace, 2001; see Beeli, Esslen, & Jänke, 2006, for the use of ERPs with exclusively sound-color synesthetes). Because ERP differences were found 200 ms after stimulus onset in more anterior regions of the scalp, rather than closer to 70 ms in posterior regions, this also points to a perceptual processing difference instead of a difference in early sensory processing stages.

It has even been suggested that the unique neurological component in synesthesia is at the attentional binding stage of processing, where color and form bind together to create a coherent representation of a stimulus. Evidence for this has been found using transcranial magnetic stimulation (TMS). Specifically, the application of a TMS current to the posterior right parietal lobe, the region of the brain involved in attentional binding, slowed the ability to match a grapheme with the synthetic color experience (Esterman, Verstynen, Ivry, & Robertson, 2006; Muggleton, Tsakanikos, Walsh, & Ward, 2007).

Recent research using diffusion tensor imaging has revealed greater white matter coherence in the superior parietal cortex, the right inferior temporal cortex, and frontal regions in synesthetes (Rouw & Scholte, 2007). Increased white matter facilitates an
increase in interconnectivity among cortical regions and provides further evidence that there are structural differences in the brains of synesthetes. This is in agreement with the idea that synesthesia involves a crossing of perceptual modalities, as well as findings that synesthesia is often derived from the mixing of signals from adjacent cortical areas (Bargary & Mitchell, 2008).

Given the neurological differences associated with synesthesia, changes in cognitive processing may follow. Indeed, synesthesia affects some aspects of cognition. For example, some synesthetes can use their experience to aid perceptual organization and to allow the perception of apparent motion that is not observed by control participants (e.g., Ramachandran & Azoulay, 2006). That said, it should also be noted that synesthesia can cause disruptions as well. When presented with words written in color, synesthetes take longer to name the color the word is printed in if that color mismatches their synesthetic experience (Dixon, Smilek, Cudahy, & Merikle, 2000; Mattingley, Rich, Yelland, & Bradshaw, 2001; Mills, Boteler, & Oliver, 1999; Nikolie, Lichti, & Singer, 2007). Moreover, Smilek, Dixon, Cudahy, and Merikle (2001; see also Smilek, Dixon, & Merikle, 2003) found that when digits were presented on a background color that matched the photisms for a digit, performance was worse, relative to when it appears on another background. Thus, the synesthetic and perceptual color processing streams appear to compete for the same neural resources. Finally, synesthetes have more vivid visual images (Barnett & Newell, 2008), which supports the idea that perceptual processing is enhanced overall, even when what is being perceived is internally generated.

**Synesthesia and Memory**

The focus of the current study is to evaluate the influence of synesthesia on memory. There are relatively few studies of synesthesia and memory, some of which are a little more than anecdotal. For example, Luria’s (1968) famous subject, S., was reported to have a phenomenal verbatim memory for information. One possible explanation is that S. used his photisms as a further memory aid. Also, although Ward (2008) highlights a significant impact of synesthesia on memory, he only lists two anecdotal accounts. Overall, it appears that synesthetes themselves tend to report that they have better memories than most people (Yaro & Ward, 2007).

Apart from the anecdotal reports, there are a few recent systematic studies of how memory is affected by synesthesia. One study by Smilek, Dixon, Cudahy, and Merikle (2002) had a single color-digit synesthete who was presented with 50-item arrays of digits. These digits were either presented in black, in colors that were congruent with the synesthete’s experience, or in colors that were incongruent with the experience. Although the synesthete performed better than most or all of the control participants for the digits presented in black or congruent colors, her performance was actually worse than that of the control participants when the colors of the digits were incongruent with her experience. Thus, this is an example of how synesthesia can disrupt memory. It should be noted that Yaro and Ward (2007) and Rothen and Meier (2009) were not able to replicate Smilek et al.’s finding using a similar paradigm. The only known replication was another single participant study presented at a conference (Azoulay, Hubbard, & Ramachandran, 2005).

Another study by Mills, Innis, Westendorf, Owsianiecki, and McDonald (2006) assessed a single grapheme-color synesthete for her memory of names, relative to a number of control individuals. Although her verbal performance was superior to the control participants, her memory for the Rey–Osterieth Complex Figure Test (Spreen & Strauss, 1998) and the Benton Visual Retention Test (Benton, 1974), both measures of spatial memory, did not differ from the control participants. A limitation of this study, as with Smilek et al.’s (2002) study, is that there was only a single synesthete. Moreover, this superior memory was assessed using only one type of material: names. Still, Mills et al. were able to show that superior memory did not generalize to information that did not elicit the synesthetic experience, suggesting that there is not just an overall memory benefit that coexists with synesthesia.

A study by Yaro and Ward (2007) found that 46 synesthetes did better on the Rey Auditory Verbal Learning Test (Spreen & Strauss, 1998) in which people were read a series of nouns, and then later tried to recall as many as they could, in any order (see also Mills et al., 2006). However, this is not a standard cognitive test of short-term memory in that it did not control for output order.

Finally, Yaro and Ward (2007) also had people complete the Rey–Osterieth Complex Figure Test in which they had to recall a complex abstract figure. As in Mills et al.’s (2006) study, there was no difference between the synesthetes and the control participants. A possible explanation for these findings was that this task did not take advantage of the synesthetic experience that occurred for verbal information. Although they were not able to replicate Smilek et al.’s (2002) finding, they were able to show that synesthetes had better memory for previously viewed color patches.

**Current Study**

We had a number of aims for the current study. First, we sought to replicate one of the more interesting findings on synesthesia and memory, namely the decrease in performance for items in inconsistent colors (Smilek et al., 2002), which others have failed to replicate (Rothen & Meier, 2009; Yaro & Ward, 2007). Second, we sought to explore the influence of synesthesia on memory beyond what has already been reported. Specifically, we were interested in situations, other than the incongruent color effect, in which additional color experiences may influence memory.

Consistent with previous research, we expected to observe some memory differences with synesthesia, although the nature of these differences is unclear. Yaro and Ward (2007) were able to reject a multicode theory, in which photism information is used to create another, more perceptually oriented memory trace, similar to Pavio’s (1969) dual code theory. Essentially, in Yaro and Ward’s study, synesthetes showed superior memory for colors, which, they argued, should not have occurred if the improved memory performance was due to dual coding (as the color patches themselves would have elicited little in the way of a verbal code). Instead, they argued that synesthetes have a generally superior memory for color information.

Another unexplored possibility is that aspects of the synesthetic experience that influence memory operate in parallel to other cognitive processes. If so, then beyond cases when perceptual and synesthetic colors are put in competition, as in the work reported by Smilek et al. (2002), there would be little impact on basic patterns of memory performance. That is, the additional synesthetic information would be encoded into the memory trace and would be boosted overall, perhaps because of their superior memory for colors, but there would be no difference in the basic pattern.
of performance across conditions. As such, standard list learning effects in memory, such as the von Restorff isolation effect (von Restorff, 1933) and the Deese–Roediger–McDermott (DRM) false memory effect (Deese, 1959; Roediger & McDermott, 1995) would continue to be observed.

A second possibility is that the synesthetic information (color in our case) becomes part of the memory trace as a feature. Moreover, synesthetes may be able to use that feature to help coordinate the retrieval of multiple memory traces. That is, people may be able to use synesthetic color as a category or memory cue to emphasize relational information between items (McDaniel & Bugg, 2008). If so, then the superior memory performance that is observed with synesthesia may be due to the better use of relational information. As such, standard memory effects that arise from relational information processing, such as the von Restorff isolation and the DRM false memory effects, would be diminished. This is because the relational information that is more typically used to separate out singletons or to relate information together would be disrupted by an emphasis on another type of relational information (McDaniel & Bugg, 2008). If so, then one would further expect to see memory recall organized less around semantic or serial order information, and more around common synesthetic colors, such as having people recall information based on the experienced colors (e.g., recall all of the “red” words first).

Finally, it may be that the synesthetic experience leads people to emphasize those characteristics of the individual items that elicit this experience. In the present case, our synesthetes all experienced colors associated with various letters and, by extension, words that they were written in. In such synesthetes, it is common for whole words to take on, to some degree, the synesthetic processing of the first letter of a word (e.g., Simner, 2007). So, the synesthetic experience may increase item-based processing at the letter and word levels apart from semantic meaning. In essence, this would result in an emphasis on item-specific processing (McDaniel & Bugg, 2008) and a general decrease in relational processing. Memory performance overall is improved because the more unique color information allows a person to specifically target a given memory trace during retrieval. Under such circumstances, standard memory effects, such as the von Restorff isolation effect and the DRM false memory effect, which depend on relational processing, would be diminished with the increase in item-specific processing. If so, then memory recall should be less organized around serial order information, and also not organized around common synesthetic colors (which would be a relational information cue).

**Experiment 1**

An important first step was to replicate the study by Smilee et al. (2002), which Yaro and Ward (2007; who used digits rather than words) and Rothen and Meier (2009) were unable to replicate. As noted earlier, Smilee et al. reported that a synesthete performed more poorly when items were presented in a color that was inconsistent with her experienced photisms for those items but not when the items were presented in black or in a color consistent with her photisms. The idea was that the inconsistency in the two types of color information competed during processing, thereby compromising performance. By replicating this effect, we hope to bring some clarity to the literature on and to set the stage to extend the investigation of the consequences of synesthesia on memory.

**Method**

**Participants.** Ten of our participants were synesthetes (eight female) who reported experiencing colors as they read different letters. They were all students at Notre Dame, the same population from which the control participants 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 appeared 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 has also been reported in other synesthetes (Simner & Holenstein, 2007; Simner & Hubbard, 2006; Smilee et al., 2007). All of the synesthetes reported what would correspond to associative synesthesia rather than projector synesthesia (Dixon, Smilee, & 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, and 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. Performance was consistent both within the sessions (strict: range = .74–.97, M = 0.90; liberal: range = .90–.99, M = 0.95) and across sessions (strict: range = .51–.91, M = 0.72; liberal: range = .71–.98, M = 0.88), 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 people were drawn from the research participant pool in the Department of Psychology at the University of Notre Dame. None of them reported having any synesthetic experiences.

**Materials and procedure.** For this task, we used lists of words because our synesthetes reported reliable synesthesia to printed words. In Experiment 1, the color that a word was pre-

1 Many of the remaining mismatches for the color mapping process were adjacent colors, such as red and brown, or yellow and orange.
sented in was varied. One third of the word lists were presented in black, and the rest were presented in color. The colored fonts could be assigned in two ways. In the Congruent condition, the color of the word matched a synesthete’s reported color for words that started with that letter on the basis of their performance in the color mapping task done originally. However, for the Incongruent condition, the color of the words did not match the synesthetic experience. These inconsistent colors were randomly sampled from a pool of other colors a person experienced for other letters. Because the synesthetes experienced different colors for different letters, the colors experienced by the control participants varied depending on a synesthete they were matched up with, using the initial letter as the guide for word color. There were 12 lists in each of the three conditions, with 12 words in each list (i.e., 432 words total). The words were 5–8 letters (M = 6.1), medium frequency (M = 131 per million as per Francis & Kucera, 1982), 1–4 syllable nouns (M = 2.0), such as “activity,” “bottom,” and “children.” The same lists and ordering were used for all participants. The words were presented one at a time, for 1 s each on a computer screen. After all of the words in a list were presented, people recalled the list by typing them into the computer in any order they wished. Only one response could be entered and seen at a time. Participants clicked a button when they were done recalling so that they could move on to the next list.

**Results and Discussion**

To provide a clearer insight into the nature of our particular synesthetes, we present the performance of each individual on the tasks in the four experiments in Table 1. The recall data, shown in Table 2, were submitted to a 2 (Group) × 3 (Condition) mixed analysis of variance (ANOVA). There was a main effect of Group, F(1, 56) = 21.80, MSE = 0.062, p < .001, with synesthetes remembering more than the control participants. There was also a main effect of Condition, F(2, 112) = 9.53, MSE = 0.002, p < .001, with performance being better in the Black and Congruent color conditions than in the Incongruent color condition. Importantly, the interaction was significant, F(2, 86) = 6.63, MSE = 0.002, p = .002.

When the control group data were analyzed separately, there was a main effect of Condition, F(2, 94) = 4.06, MSE = 0.002, p = .02. Relative to the Black condition, performance was worse in the Congruent condition, F(1, 47) = 5.30, MSE = 0.002, p = .03, and in the Incongruent condition, F(1, 47) = 7.20, MSE = 0.002, p = .01, which did not differ from one another (F < 1). So, simply adding the color to a word, in this case, perceptually, does not aid memory. Thus, the memory benefit gained by the synesthetes’ experience of color is qualitatively different.

For the synesthetes, there was a main effect of Condition, F(2, 18) = 11.49, MSE = 0.001, p = .001. In contrast to the control participants, compared with the Black condition, performance was similar in the Congruent condition, F(1, 9) = 2.00, MSE = 0.002, p = .19, but was worse in the Incongruent condition, F(1, 9) = 7.20, MSE = 0.001, p = .01. Moreover, performance was worse in the Incongruent condition than in the Congruent condition, F(1, 9) = 21.99, MSE = 0.001, p = .001.\(^2\)

In contrast to the studies by Yaro and Ward (2007) and Rothen and Meier (2009), this pattern of results is more consistent with the findings by Smirlik et al. (2002) and shows that when there is a conflict with the synesthetic experience, there is a decline in performance. The primary difference was that although our synesthetes did worse relative to themselves in the Incongruent color condition, they still did better than the control participants, whereas Smirlik et al.’s synesthete did worse. The fact that this effect is not always as strong as it was originally reported may contribute to why it has not always been found. Effectively, the competition between the perceptual and synesthetic colors disrupted the ability to effectively encode and/or retrieve the target words.

**Experiment 2**

Now that we have at least partially replicated a previous finding in the literature regarding the influence of synesthesia on memory, we can now explore how this condition may impact other aspects of memory where the additional color information may play a role. One of these was the von Restorff isolation effect (von Restorff, 1933; see Hunt, 1995, for a review). This is a classic finding in research on memory that, when given a list of words, if one of the words is uniquely defined along a given dimension it will be remembered better than the others. For example, a word printed in red has a greater probability of being remembered on a later memory test if the other words are printed in black.

The question explored here is how will synesthesia influence this finding? If control participants are presented with words printed in black, except for one printed in red, then memory for the word in red will be superior to that of the other words. However, for synesthetes, in addition to the color of the printed words, there are synesthetic colors for all of the words. As such, along a color dimension, the unique color of the singleton word may be drowned out by other color experiences. In other words, the consequence of synesthesia may be to take what was essentially a mixed list condition and turn it into a pure list condition, which is known not to reveal memory advantages seen with mixed lists (such as the bizarreness, word frequency, and enactment effects; McDaniel & Bugg, 2008). If so, then the von Restorff isolation effect for the color singleton will be reduced or absent for the synesthetes.

Alternatively, it may be that the color photisms experienced by the synesthetes, while augmenting memory for the word lists overall, will not influence the von Restorff effect per se. The color photisms may be influencing performance along a separate dimension of experience. As such, the boost in memory for a color singleton will be similar for the two groups.

**Method**

**Participants.** The same participants from Experiment 1 were also tested in Experiment 2.

**Materials and procedure.** For this task, there were 24 lists. Across the lists, all of the words were 4–7 letters (M = 5.2), 1–4 syllable (M = 1.6) nouns, which had frequencies of 0–3941 (M = 147), such as “ability,” “hair,” and “object.” The unique item in a list was identified by color and occurred in a random position within each list. Specifically, all of the words in the list were

---

\(^2\) Although the current approach of using follow-up tests without a correct for alpha may be somewhat liberal, because our p values are consistently very low, it is unlikely that a Type I error has been committed.
presented in black, except for a critical word, which was presented in red. All other aspects of the experiment were identical to Experiment 1.

Results and Discussion

The data for this experiment are shown in Table 3. Although the control participants had better memory for the critical red word relative to the black words, a von Restorff isolation effect, there was no difference between these two for the synesthetes. These data were submitted to a 2 (Group) × 2 (Condition) mixed ANOVA. There was a significant main effect of Group, $F(1, 56) = 7.82, MSE = 0.028, p = .007$, with the synesthetes remembering more words than the control participants. The main effect of Condition was significant, $F(1, 56) = 7.65, MSE = 0.012, p = .008$, as was the interaction, $F(1, 56) = 4.51, MSE = 0.012, p = .04$. Simple effects tests revealed that critical red words were remembered better than the black words for the control group, $F(1, 47) = 38.58, MSE = 0.011, p < .001$, but there was no difference for the synesthetes ($F < 1$). Thus, the synesthetes did not exploit the unique perceptual color information to aid memory further. These results are consistent with the idea that the additional color-related experiences that the synesthetes have are overwhelming any influence of a unique color singleton on a word list, effectively turning a mixed-list into a pure list (McDaniel & Bugg, 2008).

### Experiment 3

The results of Experiment 2 demonstrate that when a von Restorff isolation effect is defined in terms of a color singleton, synesthetes do not exhibit the classic effect. The possibility outlined in Experiment 2 was that the synesthetic colors experienced during study increased item-specific processes, thereby overwhelming any isolation benefit for the color singletons. If this account is correct, then if a list item is defined semantically, a von Restorff isolation effect would be observed for both synesthetes and control participants. This would be because although synesthetic color may increase item-specific processing, it should not be directly related to word meanings, and so a benefit from a semantic singleton should be observed.

Finding a semantic von Restorff effect would be consistent with an alternative explanation that also takes into account the results of Experiment 1. As noted in Experiment 1, when a word is presented in a color that is incompatible with a synesthete’s elicited phos- thoms, then this produces competition, and memory is worse. Given that all of the list singletons in Experiment 2 were in red, in most cases this would be an incongruent color condition, thereby reducing any isolation effect because the competition would be counteracting the isolation benefit. So, from this view as well, a list singleton that is based on semantic meaning rather than color would elicit a classic von Restorff isolation effect.

A third idea mentioned in the introduction is that because these synesthetes are having additional color experiences based on the physical form of the words (i.e., the letters than make them up), this results in a shift in emphasis of processing on the lexical surface form of the words at the expense of the word meanings, thereby both increasing item-specific processing and decreasing relational processing. This is especially true in the case of word lists, as these words are not part of a coherent discourse where normal word meaning derivation and application would be necessary.

The aim of Experiment 3 was to adjudicate between these various accounts. Specifically, according to color diffusing and color inconsistency accounts, when a word list singleton is iden-
tified by semantic meaning rather than color, synesthetes should show a similar von Restorff isolation effect as the control group. In contrast, if the word form emphasis account is correct, then the von Restorff effect should be reduced or absent for the synesthetes relative to the control participants.

Method

Participants. Although the same 10 synesthetes were used in this experiment, a different set of 48 participants were used. This is because the decision to explore the semantic von Restorff effect was not made until the need for it was made clear by the results of Experiment 2. By this time, the original set of control participants was no longer available. As such, a new group of people were recruited from the same population (which was also the population from which the synesthetes were drawn).

Materials and procedure. In this task, the unique item was not identified by color but by meaning. Specifically, the words were drawn from common categories, using Battig and Montague’s (1969) category norms. People were given 24 lists of 15 items each. The base category items were typically the first 14 items in a given semantic category, such as words from the category “precious stones,” such as “diamond,” “ruby,” and “emerald.” The unique 15th list item was drawn from a different and distinct category (e.g., the word “hour” in a list of precious stone names). All other aspects of the experiment were identical to Experiments 1 and 2.

Results and Discussion

The data for this experiment are shown in Table 4. For this task, like Experiment 2, the control group had better memory for the critical words relative to the control words, whereas the synesthetes showed no difference. These data were submitted to a 2 (Group) × 2 (Condition) mixed ANOVA. There was a significant main effect of Group, F(1, 56) = 12.75, MSE = 0.04, p = .001, with the synesthetes remembering more than the control participants. The main effect of Condition was not significant (F < 1). The interaction was marginally significant, F(1, 56) = 3.10, MSE = 0.01, p = .08. Again, parallel to Experiment 2, the data from the two groups were analyzed separately. When this was done, a classic von Restorff isolation effect was observed for the control participants, F(1, 47) = 10.11, MSE = 0.01, p = .003, but not for the synesthetes (F < 1). This is consistent with the idea that for the synesthetes, there is an emphasis on item-specific, surface form characteristics of the words, thereby deemphasizing factors related to word meaning. This is inconsistent with accounts based on the idea that a disruption of the isolation effect would occur if the synesthetic influences were not related to the dimension along which an item was segregated from the others.

Experiment 4

Given the result of Experiment 3, that synesthetes did not show a semantic von Restorff isolation effect, it was of interest to assess whether synesthesia would have a similar influence on another task where word meanings are important. To this end, we used the DRM false memory task (Deese, 1959; Roediger & McDermott, 1995). In this task, people are presented with a list of words that are all strong semantic associates of an additional, critical unmentioned word. At the end of the list, people are asked to recall as many words as possible. The primary finding is that people often recall the critical unmentioned word as if it were on the list.

If our interpretation of Experiment 3 is correct and our synesthetes are placing a greater emphasis on the item-specific surface form characteristics of the words, and less of an emphasis on the relational information between items, then they will be less likely to show the DRM false memory effect. This would be because there would be less of an emphasis on the semantic characteristics of the word lists. Moreover, it is likely that the critical word does not match the synesthetic photisms that accompanied the actual word list.

Alternatively, if the reduced influence of semantic processing in Experiment 3 is confined to cases where there is a semantic singleton, but does not extended to cases where all of the information is strongly interrelated, then a typical DRM false memory effect would be observed here as well. In other words, the semantic similarity of all of the items in the list may serve to increase emphasis on semantic processing in general, allowing there to be more of an influence of this type of processing.

Performance on these tasks also has implications for theories of the false memory paradigm. Ever since Roediger and McDermott’s (1995) original study, there have been a wide range of theories for this effect. Some have suggested that false memories were generated through a passive, spreading activation process (e.g., Atkins & Reuter-Lorenz, 2008; Knott & Dewhurst, 2007; Marsh & Dolan, 2007; Meade, Watson, Balota, & Roediger, 2007; Roediger & McDermott, 1995; Roediger, Watson, McDermott, & Gallo, 2001). Essentially, when the words were read, they primed their associates. Because the critical words were associated with all of the list items, they received a high degree of priming and activation, and so were then falsely recalled.

Another view is that processing strategies at encoding can influence performance. For example, processing at a shallow level (e.g., focusing on phonological aspects of words) is less likely to produce false memories than processing words at a deep level (e.g., Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999). Similarly, it has also been suggested that, consistent with the accounts of distinctive item processing described in the previous section, the false memory effect is more likely to occur when relational information is emphasized and is less likely to occur when item-specific information is emphasized (e.g., Chan, McDermott, Watson, & Gallo, 2005; McCabe, Presmanes, Robertson, & Smith, 2004). In fact, a study by Arndt and Reder (2003) showed that when the individual items were visually distinctive (using font rather than color), then this dramatically reduced the false memory effect.

Finally, there has also been some suggestion that retrieval-based factors, such as source monitoring, can impact performance on this

<table>
<thead>
<tr>
<th>Group</th>
<th>Common</th>
<th>Unique</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synesthetes</td>
<td>.82 (.03)</td>
<td>.81 (.03)</td>
<td>.82</td>
</tr>
<tr>
<td>Control</td>
<td>.62 (.02)</td>
<td>.67 (.03)</td>
<td>.65</td>
</tr>
<tr>
<td>M</td>
<td>.72</td>
<td>.74</td>
<td></td>
</tr>
</tbody>
</table>

Note. Standard errors are in parentheses.
task (e.g., Dodson & Schacter, 2001; Knott & Dewhurst, 2007; Marsh & Dolan, 2007; Roediger et al., 2001). In general, these are strategies that can be adapted during the retrieval phase to maximize performance. However in the context of the current study, all of the information is from the same source. So, although synesthetes may be able to exploit the additional color features associated with the words to produce a different pattern of performance, perhaps similar to what is seen in studies involving source monitoring, this would be of a different quality.

Method

Participants. The same participants were tested in Experiment 4 as in Experiment 3.

Materials and procedure. This task was based on the procedure outlined by Roediger and McDermott (1995). This was similar to Experiments 1–3, in that people were presented with 24 lists of 15 words that were to be recalled at the end of each list. The critical manipulation was that all of the words within a list were strongly related to an unmentioned critical word. The 24 lists provided by Roediger and McDermott were used as the study lists in this task.

Results and Discussion

The data for this experiment are shown in Table 5. Consistent with what has been reported in the literature, the control group recalled a substantial number of critical false memory words, although less often than the control words. In comparison, the synesthetes recalled the critical false memory word less often than the control participants, but they recalled the control words at a higher rate than the control participants. As such, this shows that the synesthetes exhibited a smaller false memory effect. These data were submitted to a 2 (Group) × 2 (Condition) mixed ANOVA. Although the main effect of Group was not significant, \( F(1, 56) = 2.21, \text{MSE} = 0.017, p = .14 \), the main effect of Condition was, \( F(1, 56) = 159.29, \text{MSE} = 0.025, p < .001 \), as was the interaction, \( F(1, 56) = 23.78, \text{MSE} = 0.025, p < .001 \). The more interesting comparisons here are within conditions across the two groups. These analyses showed that the synesthetes recalled far more control words, \( F(1, 56) = 22.12, \text{MSE} = 0.021, p < .001 \), than the control participants, whereas the reverse was true for the critical words, \( F(1, 56) = 8.01, \text{MSE} = 0.021, p = .006 \). Thus, the synesthetes were much less susceptible to the false memory effect.

This is consistent with the idea that the synesthetic experiencing can lead people to place a greater emphasis on surface form characteristics of the words themselves, through their synesthetic experience. Because the critical word would have not had the same color characteristics as the actual words, the synesthetes would have been less likely to report it as a false memory. Synesthesia may lead people to place a greater emphasis on item-specific surface form characteristics of the words over the semantic interrelations among the words in a list. The item distinctiveness brought about by additional visual features and the pattern of data most closely follows the data reported by Arndt and Reder (2003). Synesthetes may be able to use this additional information to keep themselves from reporting words that were not actually presented on a list. Although this would be a feature of the stimuli rather than information about the source of the words, like conditions where source information is available, synesthetes can use the additional information they have available to them to keep themselves from reporting false memories in this paradigm.

This pattern of results is inconsistent with the idea that synesthesia reduces the influence of item-to-item relational processing only when there is a single semantic item, as in the case of the von Restorff isolation effect. This is also inconsistent with any theory of DRM false memories that suggests that this phenomenon is brought about by semantic priming alone. Clearly, people use multiple sources of information to guide retrieval decisions, including the additional sensory features that are elicited by the words in our synesthetes. Thus, like other lines of research, we find that when people can bring other types of information to bear on making a memory decision, whether it be source (Arndt & Reder, 2003) or synesthetic color information, the likelihood of producing false memories is reduced.

General Discussion

The results of this study demonstrate that having synesthesia can meaningfully influence memory for lists of words. Overall, there was a general memory benefit as the synesthetes remembered more words from the lists than the control participants. This is consistent with the idea that the additional synesthetic information served to boost item-specific processing. Consistent with previous studies, we were able to at least partially replicate the finding that memory is disrupted by an incongruency between the color a word was presented in and the synesthetic color experience (Smiletk et al., 2002). Specifically, our synesthetes had poorer memory for incongruent color words compared with black and congruent color words, although their performance was still better than the control participants.

Beyond this we were able to explore whether the experience of synesthesia would lead people to place a greater emphasis on shallow, surface form characteristics of words relative to more meaningful-based aspects. Consistent with this, our synesthetes did not exhibit the standard von Restorff isolation effect, regardless of whether it was defined by color- or meaning-based singletons. Furthermore, synesthetes also exhibited a substantially reduced DRM false memory effect. Although there may be other explanations for this pattern of data, the data that synesthetes are emphasizing the words themselves, through their synesthetic experience, accounts for the entire pattern of data.

Synesthesia and Output Order

To better assess how memory processing differs in synesthetes, we also assessed how synesthesia could influence the order in which information was recalled. We approached this issue in two ways: (a) assessing the match between study and recall order, and
(b) assessing the degree to which synesthetes may have organized
their recalls by synesthetic color.

First, we assessed whether the experience of synesthesia may have
increased emphasis on processing of individual items and decreased
emphasis on the ordered relations between items (McDaniel & Bugg,
2008). If so, it is expected that synesthetes would show a smaller
influence of study order on recall order compared with control
participants. To do this, we used a relative ordering measure
(Drewnowski & Murdock, 1980) to assess the degree to which the
output order of a person’s recall attempt conform to input order
during study. This is a ratio score in which, for a given recall
length, it is assessed whether any given pair of recalled items is in
the same relative order as it was during study. Note that the relative
ordering scores can vary because of list length and are generally
low here because people were not asked to recall the words in the
order they were studied.

For Experiment 1, the relative ordering scores for the control
participants were .53 (SD = .20), .54 (SD = .20), and .53
(SD = .20) for the Black, Congruent, and Incongruent conditions,
respectively. In comparison, these scores were .41 (SD = .14),
.41 (SD = .14), and .43 (SD = .14) for the synesthetes. When
these scores were submitted to a 2 (Group) × 3 (Condition) mixed
ANOVA, there was a marginally significant main effect of Group,
F(1, 56) = 3.61, MSE = .097, p = .06, with the synesthetes’
recalls being less consistent with the input order. The main effect
of condition and the interaction were not significant (both Fs < 1).
For Experiment 2, the mean score for the control participants was .15
(SD = .05), whereas it was .10 (SD = .02) for the
synesthetes, t(56) = 3.04, p = .004. For Experiment 3, the mean
score for the control participants was .25 (SD = .12), whereas it
was .17 (SD = .09) for the synesthetes, which was a marginally
significant difference, t(56) = 1.91, p = .06. For Experiment 4, the
mean for the control participants was .34 (SD = .15), whereas
it was .20 (SD = .12) for the synesthetes, t(56) = 2.82, p =
.007. Overall, this pattern of results is consistent with an increased
item-specific memory emphasis account.

Second, we assessed whether the synesthetes were using expe-
rienced color as a retrieval cue, and clustering their recalls around
such a feature, we calculated adjusted ratio of clustering (ARC)
scores (Roenker, Thompson, & Brown, 1971) for each synesthete,
using synesthetic colors to define categories. For those readers
unfamiliar with ARC scores, a score of 1 indicates perfect cate-
gorical organization, a score of 0 indicates no organization, and a
negative score indicates some systematic organization that is
counter to the theoretical organization (color in our case). The
synesthetes had a mean overall ARC score of −.013, which did not
differ from 0 (t < 1), indicating that the synesthetes were not
organizing their recalls by color.

For Experiment 1, the data were then broken down by condition, and
ARC scores were calculated for each. Similar to the overall
analysis, there was no clear evidence that the synesthetes were
organizing their recalls by color, with ARC scores of −.013, −.051,
and −.042, for the Black, Congruent, and Incongruent conditions,
respectively, t < 1; t(9) = 1.72, p = .12; and t < 1, respectively.
For Experiment 2, the mean ARC score was .02, which did not
differ from 0 (t < 1).

For Experiment 3, unlike Experiments 1 and 2, the mean ARC score was .14, which, although small, was significantly greater
than 0, t(9) = 11.48, p < .001. Before accepting the idea that the
synesthetes were using color to organize their recall output, it is
important to note that there was an important difference between
Experiment 3 and Experiments 1 and 2. Specifically, in experi-
ments 1 and 2, the lists were constructed so that no two items
began with the same letter. However, in Experiment 3, because the
words were drawn from previously published norms, this con-
straint was not used. As such, it is possible that this significant
ARC score could be due to organizing items during recall alpha-
betically, not based on color. To assess this, we analyzed the
control participants’ data by assigning them synesthetes color cat-
egories (which is essentially random because they do not actually
experience colors). What we found was that there were significant
ARC scores here as well, t(39) = 8.45, p < .001, that were
similar to those obtained with the synesthetes, making it less likely
that there is something unique about synesthesia that is influencing
the output order of these word lists during recall.

Similarly, for Experiment 4, the mean ARC score was .11,
which, although small, was significantly greater than 0, t(9) =
7.64, p < .001. Also, like Experiment 3, because the word lists
were previously published materials, the first letter of the list items
was not controlled for. Again this significant ARC score could be
due to organizing items alphabetically during recall, not based on
color. When the control participants’ data were analyzed using the
same method as in Experiment 3, there were significant ARC
scores here as well, t(39) = 6.92, p < .001, again similar to
those obtained with the synesthetes, making it less likely that there
is something unique about synesthesia that is influencing
the output order of these word lists during recall. Overall, this pattern
of results is consistent with the idea that synesthesia increases
item-specific processing, not relational processing.

Concerns About Superior Performance

There are a couple of concerns that may be raised about the data
from this study. One is that it appears that our synesthetes have
superior memory for everything, and that this is influencing the
pattern of results. In addition to other studies that have shown this
to not be the case (e.g., Mills et al., 2006), we have other as yet
unpublished data with these same synesthetes. Overall, these data
suggest that whereas the synesthetes outperform control partici-
pants on tasks that involve the synesthetic experience, such as
memory for letters or words, there is no benefit for tasks that are
not emphasizing information that is not tied to the experience
of synesthesia, such as memory for digits, spatial processing tasks,
such as mental rotation, and language comprehension. Thus, it is
not the case that the synesthetes in our study have overall superior
memory for any kind of information, but only those types of
information that elicit the photisms.

Another concern is that some of the patterns of data, particularly
the absent von Restorff and DRM false memory effects, could be
due to the generally high rate of responding. That is, if the
synesthetes are performing at or near ceiling, it could be that this
makes it more difficult for us to detect these effects, thereby
masking the presence of the cognitive processes that typically
produce these effects. We addressed this concern in two ways.
First, we did a median split on the control participants’ control
items data and assessed whether there were any differences in
these effects in the high and low performers. Second, we selected
out those control participants that were in the same range of overall
performance as the synesthetes and assessed whether the differences were still present.

For the median split analysis, there was no meaningful influence of overall performance level on the pattern of accuracy rates. For the color von Restorff task, the low performing group people had better memory for the color singleton \((M = 0.42)\) than the control items \((M = 0.32)\), and a similar pattern was observed for the high performing group, with better performance on the color singleton \((M = 0.64)\) than the control items \((M = 0.48)\). Statistical tests revealed main effects of Group, \(F(1, 46) = 66.94, \text{MSE} = 0.013, p < .001\), and Condition, \(F(1, 46) = 39.32, \text{MSE} = 0.011, p < .001\), but the interaction did not reach significance, \(F(1, 46) = 1.89, \text{MSE} = 0.011, p = .18\). Moreover, looking at the pattern of means, although not significant, there is an increase in the size of the effect with better overall performance, not a decrease, although the synesthetes show the opposite effect.

Likewise, for the semantic von Restorff task, the low performing group people had better memory for the semantic singleton \((M = 0.53)\) than the control items \((M = 0.51)\), and a similar pattern was observed for the high performing group, with better performance on the semantic singleton \((M = 0.80)\) than the control items \((M = 0.72)\). Statistical tests revealed main effects of Group, \(F(1, 46) = 93.99, \text{MSE} = 0.015, p < .001\), and Condition, \(F(1, 46) = 10.69, \text{MSE} = 0.005, p = .002\), and the interaction was marginally significant, \(F(1, 46) = 3.68, \text{MSE} = 0.005, p = .06\), with the size of the effect being greater with better overall performance.

Finally, for the false memory task, the low performing group people had better memory for the actual items \((M = 0.45)\) than the critical items \((M = 0.15)\), and a similar pattern was observed for the high performing group, with better performance on the actual items \((M = 0.60)\) than the critical items \((M = 0.30)\). Statistical tests revealed main effects of Group, \(F(1, 46) = 79.21, \text{MSE} = 0.008, p < .001\), and Condition, \(F(1, 46) = 75.62, \text{MSE} = 0.028, p < .00\), and the interaction was not significant \((F < 1)\). So, overall, these data argue against the concern that the performance of the synesthetes is due to their overall high level of accuracy.

For the restricted range analysis, we reanalyzed the von Restorff and DRM false memory data using only those control participants that were within the same range of performance as the synesthetes. That is, they performed as well as or better than the minimum synesthetes’ performance. For the von Restorff task, there were 19 control participants that were in the same range as the synesthetes. This subset of control participants showed the standard effect, with memory performance being better on the color singleton \((M = 0.67)\) than the control items \((M = 0.49)\). There was no main effect of Group \((F < 1)\), but there was a main effect of Condition, \(F(1, 27) = 7.80, \text{MSE} = 0.017, p = .009\), and, importantly, a significant interaction, \(F(1, 27) = 5.28, \text{MSE} = 0.017, p = .03\). Thus, the same pattern of results was observed as with the entire data set.

Similarly, for the semantic von Restorff data, again there were 19 control participants that performed within the range of the synesthetes. This subset of control participants showed the standard effect, with memory performance being better on the semantic singleton \((M = 0.84)\) than the control items \((M = 0.74)\). There was no main effect of Group, \(F(1, 27) = 1.86, \text{MSE} = 0.006, p = .18\), but there was a main effect of Condition, \(F(1, 27) = 6.01, \text{MSE} = 0.004, p = .02\), and, importantly, a significant interaction, \(F(1, 27) = 10.38, \text{MSE} = 0.004, p = .003\). Again, the synesthetes did not show a semantic von Restorff effect, whereas the control participants did.

Finally, for the DRM false memory effect, for our first attempt, there was still a difference in overall performance between 25 control participants and synesthetes, \((t(33) = 3.14, p = .004)\). Looking at the data, this was likely due to one synesthete whose performance on this task was markedly different from the rest. As such, we reedited this analysis dropping this one person. In this analysis there were nine control participants that were within the range of the nine synesthetes. For this subset of control participants, memory performance was better on the actual items \((M = 0.74)\) than the control items \((M = 0.15)\). For the nine synesthetes memory performance was better on the actual items \((M = 0.79)\) than the control items \((M = 0.08)\). There was no main effect of Group \((F < 1)\), but there was a main effect of Condition, \(F(1, 16) = 919.71, \text{MSE} = 0.004, p < .01\), and, importantly, a significant interaction, \(F(1, 16) = 6.18, \text{MSE} = 0.004, p = .02\). Again, the synesthetes showed substantially smaller false memory effect than the control participants.

It might be possible to argue that there is a mirror effect (Glanzer & Adams, 1985) in this analysis, as the synesthetes had a higher hit and a lower false alarm rate than the control participants. However, if increased performance on the actual items is driving down the false alarm rate on the critical items, then a similar pattern should have been observed in the median split data, but it is not. This inconsistency makes it difficult to garner strong support for a mirror effect account of the pattern of data in the synesthetes. Instead, an account based on the experience of synesthesia affecting how memory is used is preferred. So, overall, there is no support for the idea that the absence of the von Restorff and DRM false memory effects were due to the overall superior performance of the synesthetes.

**Implications**

The results of this study have implications for theories of memory beyond our understanding of synesthesia. These data provide insight into theories of memory for unusual items (McDaniel & Bugg, 2008), such as the von Restorff effect. For the synesthetes, item specific information was boosted by the additional color information bound to the word memories. This boosted overall retrieval. However, in conjunction with an apparent influence on word graphemes, this essentially turned the task from a mixed list condition into what would be considered a pure list condition, and pure list designs often do not show influences of unusual information.

Finally, these data provide some insight into the mechanisms that can produce false memories. Specifically, the finding of a dramatically reduced false memory effect for the synesthetes argues against any account that relies exclusively on a passive spreading activation. Instead, the data are more consistent with accounts that suggest that false memories are more or less likely to be generated as a function of the strategies employed at encoding and retrieval. The reduction in the false memory effect with synesthesia lends itself to two prominent possibilities. The first is that it may be that additional color experiences with synesthesia may lead people to engage in shallower processing, focusing on the words themselves, compared with deeper level processes that can lead to false memories (e.g., Thapar & McDermott, 2001; Toglia et al., 1999). This would be consistent with our von Restorff isolation effect data.

The second is that the synesthetic experiences may increase item-specific processing because the grapheme colors serve to
highlight aspects of the individual words and are unlikely to be shared with most of the other items on a list (e.g., Chan et al., 2005; McCabe et al., 2004). Although this may also contribute to the reduction in the false memory effect, along with the shallower processing, it does not extend to and explain the absence of a von Restorff effect for the synestheses.

Overall, these data show that there are some clear influences of synesthesia on memory. However, the extent of this influence is still unclear. For example, it is not clear to what degree any memory influences at the word level scale up to larger units of information, such as sentences or even entire discourses. Also, although we feel that we have convincingly replicated Smiley et al.’s (2002) finding that a mismatch between perceptual and synesthetic color can impair memory, it is not clear whether this would extend to memory for the colors that the words are printed in as well, and even whether this extends to memory for other perceptual qualities, such as the font a word is printed in.3

Conclusions

In conclusion, the results of our four experiments show that synesthesia can affect memory. In our case, we assessed memory for word lists. Overall, our synesthetes remembered more words from any of these lists than our control participants, thus demonstrating a global benefit of memory. Also, a pattern that is apparent if one looks across all four experiments is that synesthetes show an increased use of superficial, surface form information in terms of basing memory performance more on the verbatim, graphemic characteristics of the words in the study lists, and they showed less of an effect of the semantic interrelations of the words on the list. An important issue that will need to be addressed is how synesthesia affects memory beyond word lists when memory is assessed for more complex and meaningfully interrelated materials, such as narratives and events.

References


Received March 22, 2010

Revision received August 12, 2010

Accepted August 16, 2010