2004

Eye and Mind's Eye: Evidence for Perceptually-Grounded Mental Imagery

Mark Aveyard

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

EYE AND MIND'S EYE:

EVIDENCE FOR PERCEPTUALLY-GROUNDED MENTAL IMAGERY

By

MARK AVEYARD

A Thesis submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Summer Semester, 2004

The members of the Committee approve the thesis of Mark Aveyard defended on April 27, 2004.

Rolf Zwaan
Professor Directing Thesis

Mike Kaschak
Committee Member

Chris Schatschneider
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.

“It is cold in the scriptorium, my thumb aches. I leave this manuscript, I do not know for whom; I no longer know what it is about: stat rosa pristina nomine, nomina nude tenemus.”

—*The Name of the Rose*

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

This thesis aims to show a relationship between mental imagery and sensory perception through texts that prime change blindness between mental images and visual displays. In Experiment 1, participants read short texts depicting a visual scene. Following the text, one of three types of photographs was presented: a photograph representing the scene exactly (match), a photograph representing the scene with a change (mismatch), or a photograph representing an unrelated scene (filler). Participants judged whether the picture matched the preceding text. Three types of changes were presented: color, deletion, and addition. Mirroring results from studies in visual perception, subjects showed change blindness in comparing mental images with pictures, even though critical information from the original text was recognized with a high degree of accuracy. A text variable was also included to measure the effects of language and attentional focus on change detection: each text contained a final sentence with relevant or irrelevant information related to the changed aspect of the scene. Although there were no differences in accuracy of change detection, subjects were faster to detect changes with relevant texts than irrelevant texts. Experiment 2 served as a control for comparing types of change blindness in a conventional flicker paradigm without mental imagery. As a whole, these results mirror those within change blindness studies investigating visual perception—suggesting that focused attention is necessary in comparing mental images and pictures and that mental imagery is analogous to its corresponding perceptual mode.

INTRODUCTION

Consider the following text:

Two blackbirds are perched on a wooden dock. Across the water in the distance, there is a white lighthouse on a stretch of land.

If you saw a picture of this scene now, would you be able to identify changed or missing parts? In recent years research in visual cognition has revealed a phenomenon known as change blindness, an inability to quickly detect changes in a scene or the surrounding environment (see Simons & Levin, 1997, for a review). The typical change blindness experiment involves the presentation of a visual scene in which a change occurs between one presentation and the next, with a visual buffer inserted between presentations. Almost invariably, participants struggle to detect changes that would seem obvious to any informed viewer. But what about scenes "in the mind's eye"? Does mental imagery operate in a similar way? When we read a text like the one above, can we detect changes between that mental image and similar perceptual image? What aspects of texts do we normally use in forming mental images? How detailed and accessible are those representations for comparison with other mental or visual images? And what is the nature of the mental images that stem from language? Recent research by Laeng and Teodorescu (2002) showed that eye movements during mental imagery reflect those that were enacted during the presentation of the original visual stimuli, indicating that some mechanisms of visual experience and mental imagery are used in the same way. This thesis aims to address these questions and issues further by showing a relationship between change blindness in strictly perceptual tasks and those that involve mental imagery instigated by language.

Change Blindness

Most research in change blindness has focused on change in the context of visual experience alone. The primary finding that emerges from such studies is that focused attention is necessary to detect changes in the environment (Rensink, O'Regan, & Clark, 1997). Changes that seem obvious to informed viewers or to participants *after* successful detection are often difficult to identify during the task and require considerable search time (especially in complex scenes or environments). The main theoretical insight of change blindness studies lies in their contribution to theories of visual cognition, which generally fall into one of two categories: picture theories and minimalist theories. Picture theories argue that people have a detailed mental representation of visual phenomena whereas minimalist theories contend that such phenomena are represented only by their most salient or immediately relevant parts. Change blindness studies indicate that our subjective impression of detail may not be as reliable as picture theories would indicate. If we represent visual scenes in detail but fail to see large and salient changes in those scenes, is there a way for picture theories to resolve this emerging tension? Or would we be better off favoring a more minimalist theory of visual cognition? This is the central issue around which change blindness findings revolve (Noe, Pessoa, & Thompson, 2000). Applied to mental imagery, these questions emerge as inquires into the nature of linguistic representations, the specificity of mental imagery, and the role of language in focusing attention on aspects of mental imagery.

In a review of change blindness methodologies for visual perception, Rensink (2002) identifies two approaches for stimuli presentation: the one-shot approach and the repeated change approach, which includes the oft-used flicker paradigm. In the one-shot approach, subjects have only one chance to identify change in each trial. In the repeated change approach, subjects are exposed to repeated presentations of the original stimulus and the subsequent changed stimulus. In the case of the flicker paradigm, a visible gap is inserted between the two stimuli that decreases subjects ability to form a coherent memory of each stimulus, thereby decreasing the likelihood of change detection.

For the present study, comparing mental imagery and visual stimuli requires using the one-shot approach with a brief gap between the two stimuli. This method is similar to the sentence-picture verification task which sparked the advancement of early propositional models of mental representation. In sentence-picture verification tasks, subjects read a sentence then see a congruent or incongruent picture, at which point they decide whether the picture matches the preceding sentence (Clark & Chase, 1972). Typically, negative responses to the sentence-picture relationship require more time than positive responses, and negation sentences ("The dots aren't red") require more time than affirmation sentences ("The dots are red"). Carpenter and Just (1975) created the "Constituent Comparison Model" to reflect the results of sentence-picture verification data under the assumption that people translate both sentences and pictures into abstract propositional structures, which then afford efficient intra-modal comparison. However, Tanenhaus et al. (1976) revealed serious theoretical problems in the model's construction, while additional research identified individual differences in processing strategies that the Constituent Comparison Model could not accommodate (MacLeod et al, 1978; Onken et al. 1985). There is also evidence that some people begin tasks using a perceptual approach to verification and adopt a propositional matching strategy only if required (Kroll and Corrigan, 1981).

Although Carpenter and Just attempted to reconcile their model with these challenges (Carpenter & Just, 1986), the entrenchment stood unresolved almost a decade later (see Roberts et al., 1994, for analysis and review). One of the major shortcomings of the standard sentence-picture verification tasks is the narrow scope of their materials. Sentence stimuli in these experiments nearly always take the form of a simple spatial text ("The star is above the plus") followed by iconic or line-based stimuli. Naturalistic sentences and scenes remain largely untested. Recent research indicates that naturalistic scenes may induce different cognitive processes than artificial scenes, even when they involve the same semantic material and cognitive demands (Braun, 2003). Methodologically and theoretically, sentence-picture verification studies offer little assistance in constructing meaningful change blindness experiments using mental imagery. Instead, they stand as important early battlegrounds in the argument between those who see mental representations as being largely based in perception and those who see it as largely based in abstract symbols divorced from particular sensory modes of input.

In addition to the above distinction between one-shot and repeated presentation approaches in change blindness studies, it will be helpful here to summarize the methodology of the present study within Rensink's other categories in order to clearly delineate its similarities and differences with previous change blindness studies: under Contingency of Change (the decoupling of change and motion detection) I use the "gap contingent" method (an interstimulus interval); under Content of Display (degree of realism ranging from "simple static figures" to "dynamic events in the world itself") I use "images of objects and scenes" (static, naturalistic scenes with a high degree of realism); under Content of Change (the kinds of changes made to stimuli) I use both *existence* and *properties* (addition of object, deletion of object, and color change); under Observer Intention ("degree to which an observer will expect a change") I use a *divide-attention approach* ("some other task is primary"); under Type of Task I use *detection* (subjects "simply respond to the presence of a change in the display"); under Type of Response (ranging from explicit to implicit to visuomotor responses) I use "semi-explicit" ("triggered by a 'feeling' that a change is occurring").

Goals of This Study

The aim of this study is to explore the relationship between mental imagery and perception using change blindness and to demonstrate a relationship between imagery and perception that suggests a common representational basis for each, grounded in perception. In Experiment 1, subjects heard a text describing a scene and then made verification judgements on a picture presented immediately after the text. For critical items, the picture either matched the text or contained one of three changes: color (an object's color changes), addition (an object is added to the picture), or deletion (an object is deleted from the picture). The final sentence in each text included information either relevant to the object involved in the change denoted or irrelevant to that object. For example, the text presented at the beginning of this thesis: "Two blackbirds are perched on a wooden dock. Across the water in the distance, there is a white lighthouse on a stretch of land," is followed by one of two types of sentences: "The weather is usually mild this time of year" (irrelevant) or "The weather is optimal for nesting this time of year" (related: addition/deletion) or "The weather is optimal for seeing ships out at sea," (related: color).

After reading the text, subjects see a picture exactly matching the text (the mental image generated by the text) or a picture with three blackbirds (addition), one blackbird (deletion), or a brown lighthouse (color). The subject's task is to respond YES or NO whether the picture matches the preceding text. If mental imagery is like visual perception, detection of change between a mental image generated by language and a visual percept should be analogous to detection of change between two visual percepts. Additionally, change detection rates should be higher and reaction times shorter when other information in the text highlights the situational relevance of the changed object, simulating the role of focused attention in perceptual change blindness studies. There is some precedent in visual change blindness studies for predicting this kind of context effect. Hollingworth and Henderson (2000) showed that the semantic properties of a change and the allocation of attention to parts of a scene can be dissociated in change detection. In one experiment, semantically inconsistent changes (a fire hydrant in a living room) were detected more quickly in a flicker paradigm than semantically consistent changes (a chair in a living room). In a follow-up experiment utilizing the one-shot methodology to eliminate eye-movement confounds, they obtained similar results, with greater accuracy of change detection on inconsistent trials. In a similar (though more subtle) way, the present study uses a final sentence to manipulate the saliency of change between mental image and picture.

The use of texts in place of a visual percept also affords an opportunity to look at different ways in which information about images can be represented. For example, given a text about a boy carrying red balloons, one could represent the information about this scene in a shallow manner with phonological or grammatical information or one could represent the information in a deep manner with semantic and perceptual information. To test these differences, a subsequent text recollection phase is added to each trial as a control. In this task, subjects must decide whether a word or phrase occurred in the text that came before the picture. For example, after the picture of the blackbirds appears and a decision is made regarding its match to the preceding text, subjects must then decide whether "a white lighthouse" (color test), "Three blackbirds" (addition test), and "Across the woods" (filler test) occurred in the original text. A different set of words and phrases was created for each picture condition for each item in order to test recognition memory for critical words or phrases contradicting the changed information in the picture. In other words, if a color change occurs in the picture, a critical word or phrase containing the original color information is included in the subsequent text recollection task.

Experiment 2 provides controls for the three kinds of changes involved in Experiment 1. Each changed picture used in the first experiment is matched with its exact counterpart in a conventional flicker task. Subjects are instructed to respond YES when they detect a change between the two pictures. Previous studies indicate fairly stable detection rates for addition and deletion changes, whereas color detection appears to be more variable. In one flicker study, subjects were more likely to fail in change detection with position changes (20%) than color changes (14.6%) or presence/absence changes (7.5%) (Aginsky & Tarr, 2000). When change type was cued in the same study, detection failure was highest again in position changes (20%) and much smaller in color changes (7.5%) and presence/absence changes (9.7%). With detection success, only color produced significantly different reaction times between the cued and uncued conditions (although means and MSEs were not provided). Several studies using natural scenes indicate, however, that subjects should be slower in detecting color changes than addition or deletion changes (Monty & Coltheart, 2000). Schyns and Oliva (1994) argue that dividing a scene into distinct region is one of the first processes in its perception. Color parsing, on the other hand, is of secondary importance, involving a secondary process in scene detection. The texts used in Experiment 1, however, mention color information more than it would ordinarily be mentioned. Given the varied findings in visual change blindness studies and the introduction of a text variable here, no strong predictions can be made regarding likely differences between the three change conditions. Thus, Experiment 2 is designed to measure detection rates in a conventional flicker task.

Overall, the present study makes the following main predictions:

(1) In Experiment 1, subjects will respond YES (correctly) more often in the exact condition than NO (correctly) in each of the three change conditions combined as some degree of change blindness will occur in change conditions;

(2) In Experiment 1, subjects will be faster responding NO (correctly) in the three change conditions combined than YES (correctly) in the exact condition as less time is needed for matching all relevant aspects of the text in change conditions than in exact conditions;

(3) In Experiment 1, subjects will respond NO (correctly) more often in relevant-text change conditions than NO (correctly) in irrelevant-text change conditions as the relevant-text conditions turn their attention to objects in the picture that are more likely to change;

(4) In Experiment 1, subjects will respond correctly to word/phrase recognition more often than they respond NO (correctly) in change conditions as a shallow form of the text representation is retained despite some degree of change blindness in the visual task;

(5) In Experiment 1, subjects will be faster responding NO (correctly) in the relevant-change condition than responding NO (correctly) in the irrelevant-change condition due to the attentional advantage for relevant texts noted above.

(6) In Experiment 2, subjects will show similar response times to the same stimuli used in Experiment 1 when a conventional flicker paradigm is instituted due to underlying similarities between visual and imagistic tasks.

EXPERIMENT 1

Method

Participants

Ninety-seven undergraduate students enrolled at Florida State University in introductory psychology courses for participation credits.

Materials

Sixty-eight texts were created: thirty-two critical texts and thirty-six fillers (with an additional six practice items). In order to keep subjects from adopting a "good enough" approach in matching texts with pictures (thereby ignoring critical details in the texts), eighteen of the thirty-six filler pictures depicted scenes that were semantically related to the text. For example, the text in one item describes a field of flowers with a stone wall in the foreground. Subjects then see a field covered by snow with a stone wall in the foreground. These semantically-related fillers encourage subjects to pay attention to photographic details rather than form a "good enough" strategy. Half of the filler photos were completely unrelated to the text, thereby ensuring a sufficient number of clear NO responses in the experiment, even for subjects who often fail in change detection (i.e., subjects who might adopt a simple YES response strategy). Filler items fell into one of three categories: Exact (8), Close (10), and Clash (18). Pictures were edited using Microsoft Photoshop®.

Design and Procedure

Eight lists were created to counterbalance conditions. A given text-picture item could occur in one of eight conditions: related-exact, related-color, related-addition, related-deletion, irrelevant-exact, irrelevant-color, irrelevant-addition, irrelevant-deletion. Each participant was exposed to each condition, but each item only appeared in one condition per list. This design produced a 2 (Text: Related vs. Irrelevant) X 4 (Picture Type: Exact, Color, Deletion, Addition) X 8 (List) ANOVA, with List as the only between-subjects variable. Alpha was set at .05 and beta at .20.

The experiment was run on PCs with 19" flat-screen displays using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Subjects were instructed to listen to the text (recorded by a native speaker of American English and played over speakers next to the computer) then press a key labeled "Y" for YES (right hand) if the subsequent picture matches the text or press a key labeled "N" for NO (left hand) if it does not. They were instructed to respond to subsequent word or phrases in the recognition phase of each trial by using the same keys. The complete instructions were as follows: "In this experiment you will hear descriptions of scenes from various photographs. Your task is to decide whether a picture matches the description that came before it. For each trial, you'll hear the description first. Then a picture will appear. If the picture matches the description, press the key labelled "Y" for Yes. If the picture does NOT match the description, press the key labelled "N" for No. After you make a YES/NO decision on the picture, a word or phrase will appear on the screen in quotation marks. Your task is to decide YES or NO whether that word or phrase occured in the description you heard over the headphones." There was a 1400ms delay between the end of each audio file and the appearance of the picture, which stayed on screen until the subjects made a decision (with a 100ms floor and a 10-second ceiling). After making a judgment on the picture, subjects were prompted for the word/phrase recognition task by the phrase "Did the description include . . ." which lasted for 1000ms and disappeared, followed by the first word or phrase in the recognition task for that item. Critical items included three recognition judgments, while filler items included two.

Results

Responses longer than 10000ms and shorter than 100ms were removed in order to preserve meaningful response time data (removal affected less than 2% of the data). Accuracy scores for the remaining trials were submitted to a 2 (Modality: Image/Phrase) x 2 (Picture Type: Exact/Change) x 8 (List) analysis of variance (ANOVA). List was treated as a factor in order to increase the power of the analysis by eliminating error due to random pairings of items to conditions (Pollatsek & Well, 1995), but given its lack of theoretical interest to the data will not be discussed further.

Within Image modality (judgments on pictures), subjects were more accurate in correctly responding "YES" to Exact pictures (.87) than correctly responding 'NO' to Change pictures (.65), while recognition memory for text information related to critical objects in the pictures (Phrase modality) showed a much smaller accuracy difference between trials where an Exact (.87) or a Changed picture (.86) was shown. There was a main effect of Modality [F(1,88) = 109.38, p < .0001, MSE = .001] and Picture Type [F(1,88) = 88.95, p < .0001, MSE = .002] and

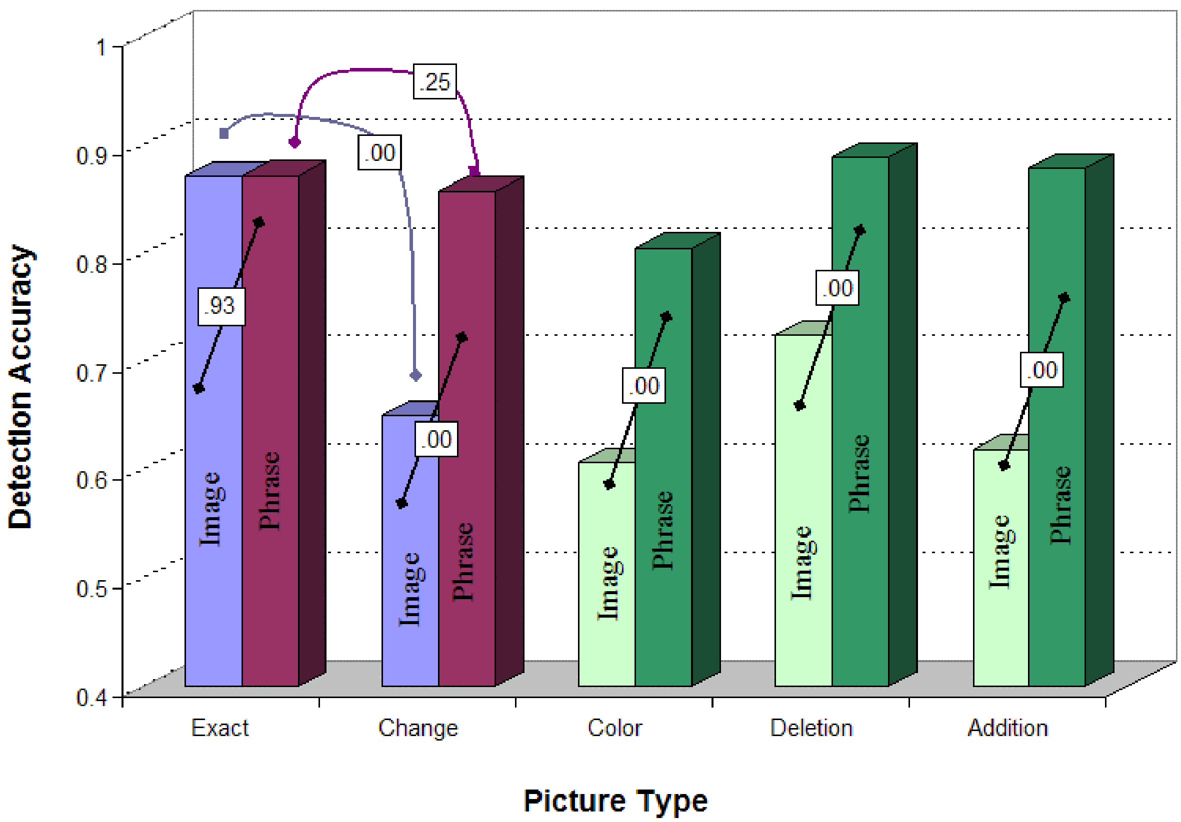


Figure 1. Change detection rates. For each pair of columns, the left column, "Image," represents change detection rates for pictures, while the right column represents change detection rates for words and phrases from the original texts. Repeated measure ANOVAs apply to "Exact" and "Change" columns. "Color," "Addition," and "Deletion" columns further specify the differences summarized in "Change". Boxed values represent t-tests for designated pairs. These values may differ slightly from reported two-factor ANOVA results due to the exlusion of list as a between subjects factor.

a reliable interaction between Modality and Picture Type [F(1,88) = 101.51, p < .0001, MSE =

.002]. This interaction was attributable to the effects of Picture Type (exact or change) on Image-Modality (pictures) [F(1,88) = 132.32, p < .0001, MSE = .004] and Modality (pictures or texts) on Change-Picture Type [F(1,88) = 181.32, p < .0001, MSE = .001], but not Picture Type (exact or change) on Modality-Phrase (word/phrase recognition) [F(1,88) = 1.53, p < .220, = MSE .001] or Modality (pictures or words) on Exact- Picture Type [F < 1]. Generally, subjects were significantly less accurate between exact and change trials in making judgments on whether the picture matched the preceding text, but there was no difference in critical phrase recognition between exact and change conditions (see Figure 1). This provides strong evidence that subjects were not merely forgetting text information whenever a change occurred in the picture but were retaining text information at similar rates across exact and change conditions.

Is lower accuracy for change compared to exact images due to the lack of focused attention on changed objects? The text variable was designed to answer this question by providing information that pertained to the changing object (Relevant) or to some other aspect of the picture (Irrelevant), thereby drawing attention toward the changed object or away from it. Thus, subjects should be more accurate and faster in identifying changes on Relevant trials than on Irrelevant trials. Accuracy scores for correct responses to Change trails were submitted to a 2 (Text: Relevant/Irrelevant) x 8 (List) ANOVA. There was no main effect for Text [F < 1]. Reaction times for correct responses to Exact and Change trails were submitted to a 2 (Text:

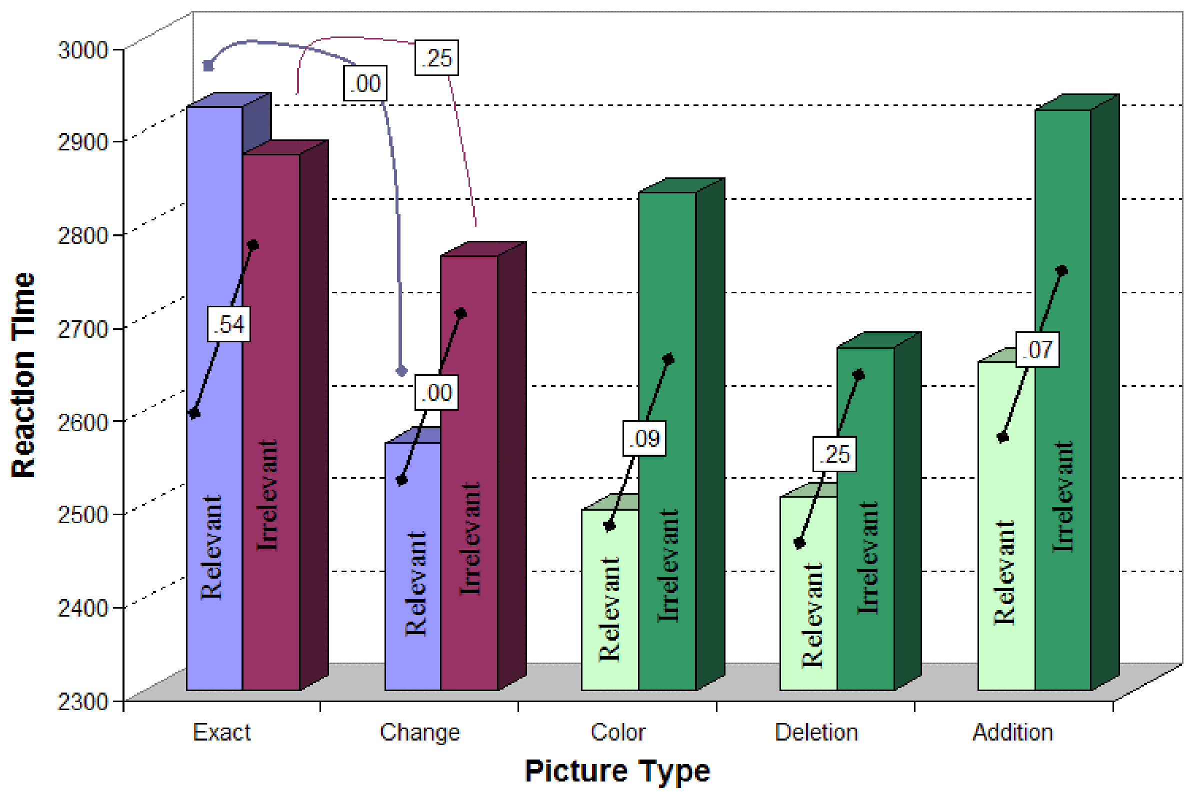


Figure 2. Reaction times for text variable (correct responses only). When attention is drawn to the changed object using semantically related information or by renaming the object, reactions times are significantly faster in correctly identifying change. Boxed values show results for t-tests.

Relevant/Irrelevant) x 2 (Picture Type: Exact/Change) x 8 (List) ANOVA. Within Text, subjects were faster correctly responding "NO" to pictures with a change (2451ms) than correctly responding "YES" to exact matches (2926ms) . As such, there was a main effect for Picture Type [$F(1,87) = 10.77$, $p < .001$, MSE = 528609.09]. There was no main for Text [$F(1,87) = 2.17$, $p < .144$, MSE 313640.91] but there was a reliable interaction between Picture Type and Text [$F(1,87) = 6.27$, $p < .014$, MSE = 231804.25]. This interaction was attributable to effects of Text (relevant or irrelevant) on Change-Picture Type [$F(1,87) = 10.79$, $p < .001$, MSE = 234285.04] and Picture Type (exact or change) on Relevant-Text [$F(1,87) = 14.58$, $p < .0001$, MSE = 380210.39] but not Picture Type (exact or change) on Irrelevant-Text [$F(1,87) = 1.61$, $p < .21$, MSE = 447189.38] or Text (relevant or irrelevant) on Exact-Picture Type [$F < 1$]. We can make two inferences from these results. First, as shown in the interaction between Picture Type and Text, altering the focus of attention in a descriptive text speeds or slows change detection. Second, although there was no main effect for Picture Type in accuracy scores, longer average response times for the Exact condition indicate that subjects engaged in a serial search comparing aspects of their mental image to the picture until they detected a change or exhausted the search criteria (see Figure 2). The absence of a main effect of text for accuracy scores indicates that there was no speed-accuracy trade-off. Moreover, the difference in reaction times between Exact and Change conditions becomes even more pronounced when collapsing the Relevant and Irrelevant text variables and looking at reaction times for correct responses regardless of text manipulation. For Exact trials, the average reaction time was 2912ms; for Change trials it was 2632ms. This generates a 2 (Picture Type: Exact, Change) by 8 (List) ANOVA with a main effect for Picture Type [$F(1,88) = 12.39$, $p < .001$, MSE = 252398.78]. This pattern was also mirrored in the reaction times for Exact and Close filler items (an example of a Close filler item would be a text that described bikers riding on a gravel road where the picture shows bikers riding on asphalt). Submitting these trials to a 2 (Picture Type: Exact and Close) x 8 (List) ANOVA shows a main effect for Picture Type [$F(1,89) = 6.37$, $p < .013$, MSE = 164318.05]. Given these results for the filler items, we cannot attribute the difference in reaction times on critical items to the type of changes being used. Rather, it appears that other kinds of change information may also be used in a serial search until a change is found in the picture.

Regarding the three types of picture changes (color, deletion, and addition), an interesting pattern of results emerges involving the change category Addition. First, addition trials showed a larger gap between Image and Phrase accuracy (.62 to .88, a difference of .26) compared to color (.61 to .81, a difference of .20) and deletion (.73 to .89, a difference of .16). Second, addition trials on correct picture responses showed longer reaction times (2782ms) than color (2610ms) and deletion (2556ms), although the difference was significant only for Addition and Deletion [$F(1,83) = 6.88$, $p < .010$, MSE = 301162.00] and not Addition and Color [$F < 1$].

EXPERIMENT 2

Method

Participants

Twenty-five undergraduate students enrolled at Florida State University in introductory psychology courses for participation credits.

Materials

Forty-four picture pairs were created: thirty-two critical pairs using materials from Experiment 1, and twelve filler items containing two additional types of changes: Location, in which an object changes location, and Switch, in which two objects change places.

Design and Procedure

Four lists matching the lists in Experiment 1 were created to counterbalance conditions. A picture pair could occur in one of three conditions: exact-color, exact-deletion, exact-addition. Each participant was exposed to each condition, but each item only appeared in one condition per list. This design produced a 3 (Picture Type: color, deletion, addition) X 4 (List) ANOVA, with List as a between-subjects variable. Alpha was set at .05 and beta at .20.

The experiment was run on PCs with 19" flat-screen displays using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Participants were instructed to find the change in each picture and press the spacebar as soon as they found it. To ensure measure detection accuracy, participants were instructed to identify the changed object by typing in a dialogue box after each trial. The full instructions were as follows: "In this experiment you will be presented with a series of flashing photographs. Your job is to detect the change that occurs in the photograph between each flash. Once you see the change, press the spacebar. A box will pop-up asking you to say what changed. Type your answer in the box." The experiment took approximately 10 minutes to complete.

Results

Accuracy rates and response times of experimental trials in were submitted to a 3 x 4 analysis of variance (ANOVA). The average reaction time for the three picture types were as follows: Color (1936ms), Deletion (1889ms), Addition (2109ms). There was no main effect for Picture Type [$F(2,42) = 1.75$, $p < .186$, MSE = 196080.62], although this may be due to lack of power. Interestingly, just as Experiment 1 showed longer reaction time for addition trials, the same picture types appear to show longer reaction times in this experiment as well. A future experiment could increase the number of subjects in order to bring this difference to significance. An additional one-shot design could also be implemented as a variation that more precisely reflects the task in Experiment 1.

GENERAL DISCUSSION

This thesis tested the hypothesis that mental imagery "in the mind's eye" operates like visual perception. Change blindness studies in visual perception show that people have surprising difficulty in detecting changes between scenes with a visual buffer. These studies also demonstrate that focused attention is necessary to successfully detect changes. In Experiment 1, subjects showed considerable change blindness when a mental image primed by a descriptive text is compared with a similar picture. Such blindness occurs despite the fact that subjects are able to subsequently recognize words or phrases from the original text contradicted by the information in the picture. Additionally, when a text variable is introduced to focus attention on objects that change or on something else, subjects were faster detecting change when the added text information was relevant to the changed object. The kinds of sentences constructed for the text variable represented a fairly subtle manipulation of attentional cues between imagery and perception. For example, in color trials 19 out of the 32 items that changed color were not mentioned by name in the final sentence; semantically relevant information was integrated through pronouns, inferences, or associated objects or behaviors ("nesting" in the case of the blackbirds).

But how exactly do we detect such changes? Do we look at everything in the picture and then make a decision? Do we engage in a serial search that terminates once a change has been detected? Do we make our comparison based on the most salient aspect of the picture? Or do we make a "good enough" judgment based on our general impression of the picture's holistic fit to the text? The reaction time data in this study indicates that we engage in a serial search by mapping aspects of our mental image onto the visual percept. In Experiment 1, subjects were consistently slower in responding correctly to Exact trials (where the picture exactly matched the text) than Change trials. Since making a correct judgment in Exact trials requires making more comparisons between the mental image and the picture, longer search time will be necessary in Exact trials compared to Change trials. Whereas in Change trials, once a mismatch has been detected, the search can be terminated. Trials in the change condition Addition were especially informative in this regard, as subjects showed longer reaction times to these trials than those of Color and Deletion. It is consistent with the general findings of the study to conclude that the presence of additional visual information related to the change requires longer search time.

The Perceptual-Propositional Divide

This thesis demonstrates a correspondence between change blindness in visual perception and change blindness in mental imagery. But does it provide any insight into the nature of mental representations? Research in mental imagery has produced an enduring divide between two primary theoretical positions: the perceptual and propositional views. Both theories use an analogical approach to description the domain of mental imagery—the perceptual view holding that it is like perception (Kosslyn, 1978; Paivio, 1978; Shepard & Chipman, 1970; Shepard and Metzler, 1971), the propositional view holding that it is like linguistic codes (Corballis, 1997; Kosslyn et al. 1997; Phelps, 1999; Farah, 1995; Chambers & Reisberg (1985). Although Experiment 1 does not provide a strong test of both views, the results do seem to at least question the plausibility of any strong version of the propositional view. Experiment 1 demonstrated that text-picture comparisons induce change blindness between mental imagery and perceptual stimuli, even when subjects are capable of recalling critical information about changed objects from the original text. Propositional theories would appear to argue that both the linguistic form of the changed objects and their perceptual manifestations are based on a common amodal code.

If this is true, it seems that a strong version of the propositional view would hold that when people fail to detect a change in a picture, they should not subsequently recognize information from the original text that pertains to the changed aspect of the picture. The results of this study show that people often do fail in the picture task but successfully recall the original text.

One objection that might arise here is that propositionally-based mental imagery could induce change blindness between imagery and perceptual stimuli due to the difficulty of "translating" or comparing the propositional code to subsequent perceptual stimuli (whereas the recognition task would involve less translational difficulty). As such, subjects might lose information temporarily during the picture verification process, especially if they feel pressured to respond quickly to the stimuli. Unfortunately there's nothing in the propositional view that *predicts* this effect, unless one *presupposes* that perceptual stimuli must be translated into propositional code before it can be used. At that point, however, the main advantage of the propositional view—its theoretical parsimony—becomes obviated by the addition of a cumbersome translation process that is no more parsimonious than the modal approach of the perceptual view. To account for these findings, the propositional view would have to hold that subjects maintain two codes simultaneously for different tasks or that the cognitive architecture employs the base code in different ways for different tasks. The perceptual view, on the other hand, argues that only one imagistic code is necessary, a modal code that behaves like its perceptual counterpart. In this way, the perceptual view places the complexity of the system in the fundamental symbols of mental imagery, describing them as modal in nature and simplifying later processes. The propositional view places the complexity of the system in its performance.

Difficulties with Mental Imagery

Although cognitive scientists have staked claims on what mental imagery is *like*, neither the perceptual view nor the proposition view attempts to explain what mental imagery *is*. Using metaphor or analogy as the basis of a scientific theory (rather than a literal descriptive approach) implicitly begs questions regarding the nature of the phenomena being explained because a metaphor necessarily describes only part of its target domain. When used as the foundation of a theory, metaphor risks presenting a woefully incomplete explanation. Even so, its use can be justified when it provides more insight than preceding viewpoints, although the limitations of this approach should be acknowledged.

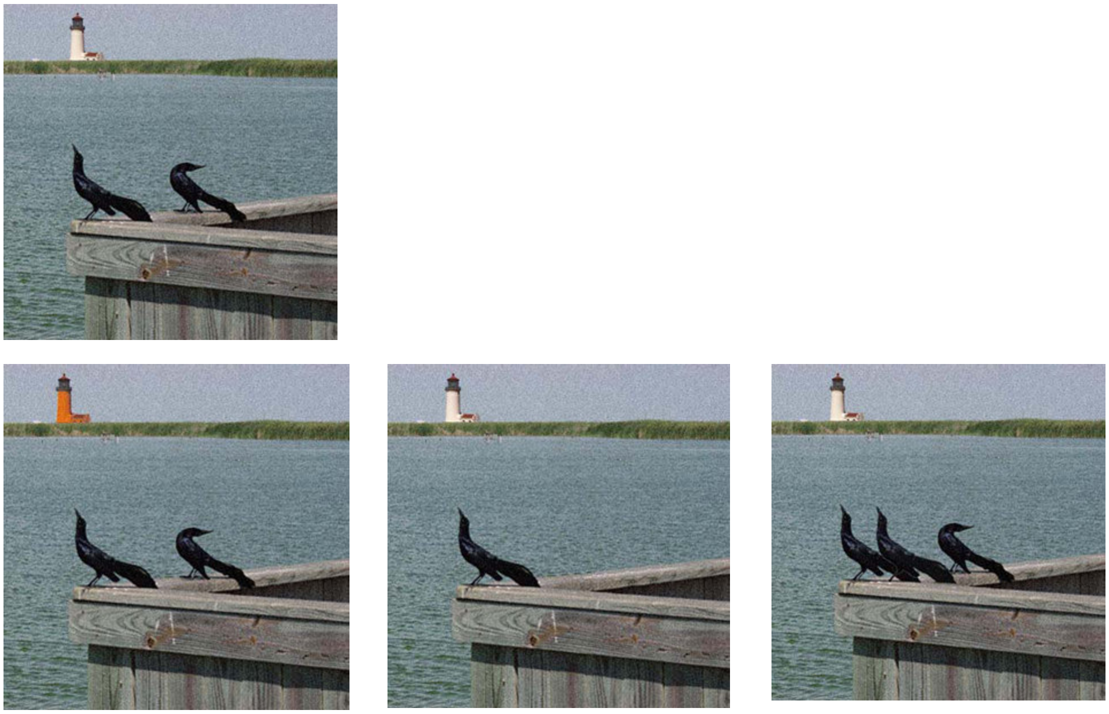
Another common pitfall in mental imagery theories is the confusion of process with identity, of confounding the *mechanisms* used in generating mental imagery with its *ontology*. Such is the error manifest in "eye movement" theories of imagery in so far as those theories *define* mental imagery as eye movements in the absence of perceptual input. When the physical mechanism can be used or executed without causing the cognitive phenomena in question—or visa versa—such identity theories prove insupportable. A person who loses his eyes is still capable of experiencing mental imagery. A congenitally blind person may have eye movements without instigating visual mental imagery (Zimler & Keenan, 1983).

An additional impetus to fruitful research and theorizing on mental imagery is the widespread mistrust of subjectivity on the part of cognitive scientists today. Recall that psychology as a professional field of research began with a nearly ubiquitous emphasis on subjective reports of mental states. Despite the more recent decline of behaviorism, in which mental states held no empirical validity, this predilection against subjectivity has carried over to contemporary cognitive psychology. Such a bias can be attributed in part to the inherent difficulties in rendering subjective reports or introspection into reliable statistical analysis (especially with small sample sizes). But the methodological difficulties presented by a phenomenon do not obviate its ontological character. As such, the current entrenchment between the perceptual and propositional schools may very well hinder an ontology of mental imagery until the possibly unsubstantiated dogma against subjectivity is curtailed or abandoned. It is interesting to note that one of the early seminal studies in mental imagery, the original mental rotation study by Kosslyn, owes its genesis to a lucid dream, featuring a splendid rotating object in the sky, experienced by that study's author one morning shortly before waking up (Matlin, 2002).

The problem of metaphor and the mistrust of subjectivity are both impediments to a sound ontology of mental imagery and to a wider range of empirical research. The purpose of this study, though, is not to precisely elucidate the nature of mental imagery but to add knowledge to the contemporary debate within the limitations of its metaphor-grounded approach. Ultimately, whether one adopts one view or the other depends partly on the emerging body of evidence in mental imagery and perceptual research and partly on considerations of the teleology of human cognition, that is, on considerations of what the fundamental 'building blocks" of mental imagery and perception are *for* in the first place. While this thesis does not address issues of teleology in cognition, it adds more evidence to a converging body of research on the mental imagery and perception that as a whole indicates the superiority of the perceptual view.

APPENDIX A

Example of Pictures Used in Experiments 1 and 2



APPENDIX B

Texts Used in Experiment 1 in Color Trials

RELEVANT	**IRRELEVANT**
Two yellow flowers are at the bottom of a sand dune that slopes up to a partly cloudy sky. It looks like a bad place for vegetation.	Two yellow flowers are at the bottom of a sand dune that slopes up to a partly cloudy sky. It looks like a place far from any town.
Two blackbirds are perched on a wooden dock. Across the water, there's a white lighthouse on a stretch of land. The weather looks good for seeing ships at sea.	Two blackbirds are perched on a wooden dock. Across the water, there's a white lighthouse on a stretch of land. The weather looks good for tourists.
There's a church tower with a red steeple on top next to a sanctuary with a clock above its door. There are no windows on the steeple tower . The tower is roughly twice a tall as the sanctuary.	There's a church tower with a red steeple on top next to a sanctuary with a clock above its door. There are no windows on the steeple tower. The grass looks somewhat dry.
Two cows are in a green field. There's a red wheelbarrow in the background and the closest cow is munching on the grass. The wheelbarrow looks empty.	Two cows are in a green field. There's a red wheelbarrow in the background and the closest cow is munching on the grass. There are little yellow flowers in the field.
A dog is jumping into a round swimming pool to fetch a red ball. There are no people or anything else in the pool. The ball is floating.	A dog is jumping into a round swimming pool to fetch a red ball. There are no people or anything else in the pool. The pool looks clean.
A blue dragonfly with four wings is on some kind of rough surface. It looks like it's resting.	A blue dragonfly with four wings is on some kind of rough surface. There's nothing else there.
There's a two-door garage adjacent to a house. The garage doors are a dark reddish-brown. The driveway is empty. The doors are closed.	There's a two-door garage adjacent to a house. The garage doors are a dark reddish-brown. The driveway is empty. There are trees behind the garage.

There's a rectangular dining table with white chairs around it. Some white placemats are the only things on the table. There's one for each chair.

There's a rectangular dining table with white chairs around it. Some white placemats are the only things on the table. There's a wood floor.

Near a large tree, two sheep are grazing on green grass. There's no one tending them.

Near a large tree, two sheep are grazing on green grass. There are some clouds in the sky.

A boy in shorts and a hat is holding onto a string of three red balloons. There are bushes in the background. It looks like the wind is blowing hard.

A boy in shorts and a hat is holding onto a string of three red balloons. There are bushes in the background. The boy is probably no more than four years old.

A big yellow umbrella in the sand is sheltering two empty beach chairs that are facing the ocean. The shade looks comforting.

A big yellow umbrella in the sand is sheltering two empty beach chairs that are facing the ocean. It looks like a peaceful day.

There are two paint brushes, an orange tape measure, and some steel wool resting on a wooden table. Someone might be getting ready to measure something.

There are two paint brushes, an orange tape measure, and some steel wool resting on a wooden table. Someone might finishing a project.

There are three blue buckets sitting on a tile floor. One of the buckets has some flowers in it. There's enough volume for several gallons of water.

There are three blue buckets sitting on a tile floor. One of the buckets has some flowers in it. The tile looks as though it's been cleaned recently.

There's an entrance to a brick office building with a door mat leading to the front doors and a large green plant holder on each side. Some large plants would look good in them.

There's an entrance to a brick office building with a door mat leading to the front doors and a large green plant holder on each side. The doors are reflective.

Three short green candles are sitting on a nightstand in front of a couple bottles of perfume. There's nothing to keep wax off the table.

Three short green candles are sitting on a nightstand in front of a couple bottles of perfume. The perfume bottles are round.

Two blue shopping carts are standing in the parking lot of a department store with a green front entrance. No one is going in or out.

Two blue shopping carts are standing in the parking lot of a department store with a green front entrance. It's a sunny day.

Two square blue coasters are resting on a black table. Behind the coasters, there's a stone white cross. There are no drinks on the table.

Two square blue coasters are resting on a black table. Behind the coasters, there's a stone white cross. The desk has a few chips.

There are two blue cones on a road. On the grass in the background, there's a row of portable toilets. The cones are evenly spaced.

There are two blue cones on a road. On the grass in the background, there's a row of portable toilets. The grass is a little dry in places.

There's a copier with two red books on top of its lid. The books are closed.

There's a copier with two red books on top of its lid. The output tray is empty.

There's a black leather couch with two red pillows and a box of kleenex sitting on one armrest. It looks long enough for sleeping.

There's a black leather couch with two red pillows and a box of kleenex sitting on one armrest. There's a rug in front of the couch.

There's a coffee machine on a table with two red cups next to it. The cups look empty.

There's a coffee machine on a table with two red cups next to it. The coffee pot looks empty.

There are two blue satellite dishes rising behind a single-story brick building. They're about the same height as the roof.

There are two blue satellite dishes rising behind a single-story brick building. There are bushes around the building.

There are two red dumbbells near a sock on the carpet. They don't look particularly heavy.

There are two red dumbbells near a sock on the carpet. There's nothing else on it.

There are two blue dumpsters in an empty parking lot. The building in the background has a slanted roof. There's no litter nearby.

There's a pole with two rectangular red banners. A couple of students with backpacks are walking by in the background. There's a light at the top of the pole.

Two green golf carts are sitting in front of a fairway. There's a pond to one side in the background. There are no clubs in the carts.

There's a knife holder on a kitchen counter. In it, there are three knives with blue handles. Some of the slots don't have knives.

There are two red mailboxes in front of a wooden fence. The ground beneath them is covered in snow. Their doors are closed.

Two parachuters with yellow chutes are floating down onto a field. There's a barn and silo at one end of the field. The chutes aren't completely round.

There are two red pay phones next to a bench. The bench is empty except for a black shoulder bag. The phones aren't being used.

There are two green plates on a wood table in front of a purse. One plate has a fork on it.

There are two blue dumpsters in an empty parking lot. The building in the background has a slanted roof. There are no cars in the parking lot.

There's a pole with two rectangular red banners. A couple of students with backpacks are walking by in the background. They're facing away from the pole.

Two green golf carts are sitting in front of a fairway. There's a pond to one side in the background. The fairways look a little brown.

There's a knife holder on a kitchen counter. In it, there are three knives with blue handles. There's an outlet on the wall behind the knives.

There are two red mailboxes in front of a wooden fence. The ground beneath them is covered in snow. There are trees in the background.

Two parachuters with yellow chutes are floating down onto a field. There's a barn and silo at one end of the field. There are some people on the field.

There are two red pay phones next to a bench. The bench is empty except for a black shoulder bag. There's a bush in the background.

There are two green plates on a wood table in front of a purse. There's a set of keys near the purse.

There are two sailboats on the water. The one in the background has a red and blue main sail. The boats are headed in the same direction.

There are two sailboats on the water. The one in the background has a red and blue main sail. There's nothing else on the water.

References

Aginsky, V., Tarr, M. J. (2000). How are different properties of a scene embodied in visual memory? *Visual Cognition, 7,* 147-162.

Altmann, G. & Kamide, Y. (2003) Embodying the mapping between language and the visual world: Dissociations between the visual world and its mental representation. *The 16th Annual CUNY Conference on Human Sentence Processing.*

Braun, J. (2003). Natural scenes upset the visual applecart. *Trends in Cognitive Sciences, 7(1),* 7-9.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 82, 45-73.

Chambers, D., & Reisberg. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception & Performance, 11*, 317-328.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472-517.

Corbalis, M. C. (1997). Mental rotation and the right hemisphere. *Brain and Language, 57,* 100-121.

Farah, M. J. (1995). Current issues in the neuropsychology of image generation. *Neuropsychologia, 33,* 1455-1471.

Grimes, J. (1996). On the failure to detect changes across saccades. Akins, Kathleen A. (Ed). (1996). *Perception: Vancouver studies in cognitive science, 5* (pp. 89-110). London, Oxford University Press.

Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of change in natural scenes. *Visual Cognition, 7,* 213-235.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception & Performance, 4,* 47-60.

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *Neuroimage, 6,* 320-334.

Kroll, J. F., Corrigan, A. (1981). Strategies in sentence-picture verification: The effect of an unexpected picture. *Journal of Verbal Learning & Verbal Behavior, 20,* 515-531.

Laeng, B., and Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26,* 207-231.

MacLeod, C. M., Hunt, E. B., Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior, 17,* 493-507.

Matlin, M. W. (2002). *Cognition.* Fort Worth: Hartcourt College.

Noe, A., Pessoa, L., Thompson, E. (2000). Beyond the grand illusion: what change blindness really teaches us about vision. *Visual Cognition, 7,* 93-106.

Onken et al. (1985). Individual differences in the use of simplification strategies in a complex decision-making task. *Journal of Experimental Psychology: Human Perception & Performance. 11,* 14-27.

Paivio, A. (1971). Comparison of mental clocks. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 61-71.

Phelps, E. A. (1999). Brain versus behavioral studies of cognition. In R. J. Sternberg (Ed.) *The nature of cognition* (pp. 295-322). Cambridge, MA: MIT Press.

Pollatsek, A., & Well, A.D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 785-794.

Rensink, R. (2002). Change Detection. *Annual Review of Psychology*, 53, 245-77.

Rensink, R. A., O' Regan, J. K., & Clark, J. (1997). To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368-373.

Rich, A., & Gillam, B. (2000). Failure to detect changes in color for lines rotating in depth: the effects of grouping and type of color change. *Vision Research, 40,* 1377-1384.

Roberts, M. J., Wood, D. J., & Gilmore, D. J. (1994). The sentence-picture verification task: methodological and theoretical difficulties. *British Journal of Psychology, 85,* 413-432.

Ruse, M. (1973). The Value of Analogical Models in Science. *Dialogue, 12,* 246-253.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh: Psychology Software Tools, Inc.

Schyns, P.G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science*, 5, 195-200.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology, 1,* 1-17.

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

Simons, D. J., Franconeri, S. L., and Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception, 29,* 1143-1154.

Simons, D. J., and Levin, D. T. (1997). Change Blindness. *Trends in Cognitive Sciences, 1*, 261-267.

Tanenhaus, M. K., Carroll, J. M., & Bever, T. G. (1976). Sentence picture verification models as theories of sentence comprehension: A critique of Carpenter & Just. *Psychological Review, 83,* 310-317.

van der Heijden, A. H. C. (1993). The role of position in object selection in vision. *Psychological Research/Psychologische Forschung, 56,* 44-58.

Zimler, J., & Keenan, J. M. (1983). Imagery in the congenitally blind: How visual are visual images?. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 9,* 269-282.

BIOGRAPHICAL SKETCH

Mark Aveyard was born November 18, 1975, in Denver, Colorado, and moved to Nebraska at the age of five. He graduated from Beatrice High School in Beatrice, Nebraska, in 1994, and from the University of Nebraska-Lincoln in 1998 with a B.A. in political science.

He became fascinated with cognitive science generally and psycholinguistics in particular after a long and on-going love affair with literature, poetry, and language ("In the beginning was the Word.") The aim of his research is to illuminate aspects of language that relate to literary experience and to the architecture of the mind that surrounds it.

> "Men do not *invent* those mysterious relations between separate external objects, and between objects and feelings or ideas, which it is the function of poetry to reveal. These relations exist independently, not indeed of Thought, but of any individual thinker. And according to whether the footsteps are echoed in primitive language or, later on, in the made metaphors of poets, we hear them after a different fashion and for different reasons. The language of primitive men reports them as direct perceptual experience. The speaker has observed a unity, and is not therefore himself conscious of *relation*. But we, in the development of consciousness, have lost the power to see this as one. Our sophistication, like Odin's, has cost us an eye; and now it is the language of the poets, in so far as they create true metaphors, which must *restore* this unity conceptually, after it has been lost from perception. Thus, the 'before-unapprehended' relationships of which Shelley spoke, are in a sense 'forgotten' relationships. For though they were never yet apprehended, they were at one time seen. And imagination can see them again."
>
> —Owen Barfield