False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness blindness

Daniel T. Levin,* Sarah B. Drivdahl, Nausheen Momen, and Melissa R. Beck

Department of Psychology, Kent State University, P.O. Box 5190, Kent, OH 44242-0001, USA

Received 3 July 2002

Abstract

Recently, a number of experiments have emphasized the degree to which subjects fail to detect large changes in visual scenes. This finding, referred to as “change blindness,” is often considered surprising because many people have the intuition that such changes should be easy to detect. Levin, Momen, Drivdahl, and Simons (2000) documented this intuition by showing that the majority of subjects believe they would notice changes that are actually very rarely detected. Thus subjects exhibit a metacognitive error we refer to as “change blindness blindness.” Here, we test whether CBB is caused by a misestimation of the perceptual experience associated with visual changes and show that it persists even when the pre- and postchange views are separated by long delays. In addition, subjects overestimate their change detection ability both when the relevant changes are illustrated by still pictures, and when they are illustrated using videos showing the changes occurring in real time. We conclude that CBB is a robust phenomenon that cannot be accounted for by failure to understand the specific perceptual experience associated with a change.

© 2002 Elsevier Science (USA). All rights reserved.

* Corresponding author. Fax: 1-330-672-3786.
E-mail address: dlevin@kent.edu (D.T. Levin).
1. Introduction

Recently, a number of researchers have found that subjects are surprisingly poor at detecting changes in visual scenes. These omissions occur both in complex natural scenes, and in artificial displays whether or not subjects are expecting the change (see Rensink, 2002; Simons & Levin, 1997, for reviews). These findings challenge psychological theories about the representation of visual detail, and the specificity with which attention needs to be directed in order to detect unexpected visual events. In addition, change blindness (CB) often comes as a surprise to students and the public who seem to believe that they would have no problem detecting such changes. We consider this latter observation to be important because it implies that people have a set of beliefs about their visual system that may not match its true capacities. This metacognitive error was confirmed empirically by Levin, Momen, Drivdahl, and Simons (2000), who found that a large percentage of subjects predicted they would detect changes that few subjects in actual experiments noticed. We refer to this finding as “change blindness blindness” (CBB). Here, we test whether CBB occurs because subjects falsely believe that a perceptual experience akin to pop-out will accompany changes. Contrary to this hypothesis, we show that CBB persists when the nature of the change event precludes this experience, and when the actual stimulus films used in Levin and Simons (1997) are used as the basis for subjects’ estimates. Based on these findings we argue that it CBB is enhanced by a variety of factors including beliefs in identity tracking, and a misapprehension of the relationship between attention and scene perception.

1.1. Change blindness and change blindness blindness

Although much research has explored metacognitive issues in memory, very little is known about how people understand their visual system (for some interesting exceptions see Flavell, Green, & Flavell, 1995; Miller & Weiss, 1982; Winer, Cottrell, Karefilaki, & Chronister, 1996; Winer, Cottrell, Karefilaki, & Gregg, 1996). Despite the lack of formal research on adult visual metacognition, a number of findings basic to vision and visual memory strike psychologists and lay observers as counterintuitive. In particular, recent research has emphasized situations where the visual system is surprisingly ineffective. For example, Rensink, O’Regan, and Clark (1997) showed subjects two rapidly alternating versions of a scene. Subjects often required many alternations to see the difference between the scenes even though it was obvious once spotted. This, and a variety of related findings all support the conclusion that people do not automatically retain and compare visual details between views (for example, Blackmore, Brelstaff, Nelson, & Troscianko, 1995; Grimes, 1996; Henderson, 1997; McConkie & Currie, 1996; Simons, 1996; for a review see Simons & Levin, 1997). Instead, it appears necessary to pay attention to the specific thing that changes in order to see the change. This is not, however, to say that attending to an object is sufficient to see it change. In Levin and Simons (1997), subjects viewed motion pictures in which the sole actor suddenly changed identity across the cut between shots. Even though this change occurred in the central object in the scene, most subjects failed to notice it. Simons and Levin (1998) extended this finding by
demonstrating that approximately 50% of subjects also miss the substitution of their real-world conversation partner.

Change blindness is interesting for a number of reasons having to do with attention and visual memory, but it is also arresting because it conflicts so strongly with intuition. In fact, prior to running these experiments, many people, ourselves included, were convinced that the changes would be easily noticed. Levin et al. (2000) confirmed this conflict by asking subjects if they would see changes that usually went unnoticed by the subjects in Levin and Simons (1997) and Simons and Levin (1998). In most cases, the vast majority of subjects we tested claimed they would detect these changes, and were quite confident in their responses. For example, Levin and Simons (1997) found that 0% of subjects noticed the disappearance of an actor’s scarf while 90% of subjects in Levin et al. (2000) claimed they would have seen this change.

It is important to emphasize from the outset that CBB is not a global overconfidence effect (Fischhoff, 1988; Metcalfe, 1998). First, this kind of general overconfidence effect is generally relatively small (about 20%) relative to CBB, and Levin et al. (2000) found that CBB was not affected by changing the scenarios to make them estimates of other people’s performance (e.g., subjects predicted if “someone” would notice the change instead of predicting their own performance). In addition, the same subjects who overestimate change detection ability also underestimate their ability to recognize a large number of pictures (Levin, 2001).

If CBB is not simply an overconfidence effect, then what causes it? In this report we test whether CBB occurs because subjects believe that changes will call attention to themselves via a perceptual transient comparable to apparent motion. Accordingly, they would not be far from the truth because change blindness usually does depend on some sort of masking to eliminate this transient (see Rensink, 2002; Simons & Levin, 1997, for reviews). Therefore, the experiments reported in this paper include conditions that de-emphasized possible transients by asking subjects to judge if they would detect a change in which the prechange view was separated in time from the postchange view. This manipulation might be expected to eliminate CBB because it eliminates the presumed transient, and also should make clear that change-detection requires some sort of visual memory for detail whereby information from the prechange view must be retained and compared with the postchange view. Thus these studies may prompt subjects to recruit their understanding of memory to correctly answer change-detection questions (e.g., Flavell, Friedrichs, & Hoyt, 1970; Miller & Weiss, 1982; Tversky, 1973).

In addition, we explored the conceptual basis of CBB by asking subjects open-ended questions about beliefs relevant to the CBB scenarios. In Experiment 2, we asked subjects what factors affected their responses in each scenario, and coded their responses for mention of a number of factors including pre- and postchange similarity, the roles of memory and attention, and object centrality. Finally, in Experiment 2 subjects made estimates based on videos instead of still images. These videos consisted of the actual stimuli used in Levin and Simons (1997) or depicted a subject’s-eye view of the real-world person change in Simons and Levin (1998).
2. Experiment 1

In Experiment 1 we tested subjects’ beliefs about their change-detection ability and their digit span. We used the same five scenarios used in Levin et al. (2000). Of the four change-detection scenarios, two involved changes to objects, one of which was distant from the putative focus of, and one of which was closer to the center of attention. These will be referred to as object changes. The other two changes, which will be referred to as identity changes, were substitutions of one person for another in the center of attention.

All estimates were tested under three different conditions. In the first (the immediate condition), all scenarios were similar to those in Levin, Momen, Drivdahl, and Simons (1999) in that they referred to unexpected changes that occurred over very short time spans. In the second and third conditions (the minute and hour conditions), the same scenarios included a 1-min or 1-h delay between the pre- and postchange views (or between exposure to the digits and the memory test). Experiment 1 therefore tested whether CBB would disappear or be reduced when conditions preclude a perceptual transient, and need to retain visual detail is made salient by delays separating pre- and postchange views.

In addition, subjects completed a postexperiment questionnaire in which they indicated whether they thought it is necessary to pay attention directly to the changing object in order to detect the change. This question was added to assess the degree to which CBB was related to beliefs about attention based on previous research documenting adults’ understanding that memory is improved when one attends to the relevant object (Miller & Weiss, 1982).

2.1. Method

Participants. Fifty-one Kent State University undergraduates participated in this experiment (13 in the immediate condition, 20 in the minute delay condition, and 18 in the hour delay condition) in exchange for credit in their general psychology course.

Materials. Four scenarios describing unexpected between-view changes similar to those used in Levin et al. (2000) were created. Each described a specific situation tested by Levin and Simons (1997) or Simons and Levin (1998), who found that surprisingly few subjects actually detected the changes. In scenarios 1, 2, and 3 subjects were asked to imagine that they were watching a movie in which some visual feature unexpectedly changed across views. In scenario 1 (plate change), the movie shows two actors having a conversation in a restaurant. In the middle of the conversation, the plates on their table change from red to white. In scenario 2 (scarf change), a scarf one of these actors is wearing disappears across a cut. Levin and Simons (1997) found that 0% of subjects actually noticed these changes. Together, these two scenarios will be referred to as “object-change” scenarios. Scenario 3 (actor change) describes a movie in which a single actor can be seen sitting in his office. He looks up and walks into the hall to answer a telephone, but when the view switches from the office to the hall, the first actor has been replaced with a second actor (wearing different clothes) across the cut.
Again, Levin and Simons (1997) found that 0% of subjects actually noticed the change. Scenario 4 (pedestrian change) describes a real-world instance of change blindness in which the subject imagines being approached on the street by a lost pedestrian asking for directions. In the middle of the conversation, two people carrying a door momentarily interrupt the conversation by rudely walking between the subject and the pedestrian. When they have passed, a different person is left behind to continue the conversation. Simons and Levin (1998) found that approximately 46% of subjects noticed this change. The actor and pedestrian scenarios will be referred to as “identity-change” scenarios. All change scenarios ended with a reminder that “You did not expect this change beforehand.”

Each of the four change blindness scenarios was accompanied by color illustrations showing the pre- and postchange views. For scenarios 1–3 these were digitized stills from the actual stimulus videos, and scenario 4 was illustrated by a side-by-side still of the two experimenters from Simons and Levin’s (1998) Experiment 2. The last scenario described a typical digit span test in which the subject heard an experimenter read a series of digits which they were to repeat back correctly.

Scenarios in the delay conditions described a delay between the pre- and postchange view. For the movie scenarios (plate change, scarf change, actor change) subjects were asked to imagine pausing the VCR, and getting up to answer the phone between the pre- and postchange views views. For the pedestrian change scenario, subjects were asked to imagine that they had a conversation with a friend for 1 min or 1 h while the people carrying the door blocked their view of the lost pedestrian. All delays were filled with verbal behavior and new visual experience. For example, the movie scenarios specified that the phone was “in another room” which we hoped would emphasize that a new set of visual experiences would intervene between the pre- and postchange view. In the digit span scenario subjects were asked to imagine that the delay was filled with a subtraction task (e.g., starting at 1000 and counting backwards by sevens).

The scenarios and accompanying illustrations were assembled into packets. The first page of each packet gave an introduction to the experiment which described the process of combining views, and a brief description of the necessity to retain visual information across views in order to detect changes which included explicit reference to the possibility that “we use our visual memory to combine glances.” The introduction also included a reminder that the subject should imagine that he/she was not on the lookout for changes. The change blindness scenarios followed the cover page in pseudo random order. The text for each scenario was printed on one sheet, and the illustration on the following sheet. The digit span scenario was printed on the last sheet of the packet.

An open-ended questionnaire asking subjects about their responses was also added. It included the questions: “In order to see these unexpected changes, do you think it is necessary to pay attention directly to the object that changes, or would you just notice the change without paying attention to that particular object?” and “Is the above response based on any particular experiences you have had? If so, what?” Subjects also indicated whether they had heard about change detection research,
research testing the number of digits people can remember, or the “magic number seven.”

Procedure. Subjects were run in small groups ranging in size from 1 to 6. They were each given a packet and told to look at the cover sheet while it was read to them by the experimenter. Subjects were then instructed to go through the packet by first reading each scenario, then looking at the picture illustrating the change. For each scenario, subjects judged whether they would notice the change, and indicated how confident they were in their prediction using a 7-point scale ranging from “not very confident” (rating of 1) to “very confident” (rating of 7). For the digit span scenario, subjects judged whether they could remember 10 digits.

2.2. Results

Change-detection scenarios. As in Levin et al. (2000), subjects overestimated their change detection ability (see Table 1). Based on the empirical base rates established in Levin and Simons (1997) and Simons and Levin (1998), subjects could be expected to notice an average of 11.5% of the four changes. Mean estimates were significantly higher in all conditions. Subjects indicated they would detect 75% of changes in the immediate condition, 72.3% in the minute condition, and 64.8% of changes in the hour condition, $F(2, 48) = 1.418$, $MSE = .630$, $p = .2521$; see Table 1. A planned contrast testing the difference between the immediate condition and the hour condition was also nonsignificant, $F(2, 48) = 2.366$, $MSE = .630$, $p = .131$. Estimates were not significantly less in the hour condition than in the minute condition in any of the four individual scenarios, although the plate scenario came close, $X^2 = 3.53$, $p < .10$; see Table 1. Estimates for all scenarios in all conditions were again significantly greater than their immediate empirical base rates ($p < .05$, Fisher’s test).

---

1 Although some subjects had heard about change blindness, and about half reported knowledge about short term memory, their responses to these questions were not related to change blindness or responses on the digit scenario in either Experiments 1 or 2. In Experiment 1, A total of 32 subjects claimed to have heard about the capacity of short term memory, and 18 knew the magic number seven. Four of the eight subjects in the immediate condition who responded positively to the short term memory question indicated they could remember 10 digits, while one of the four who had not heard about this predicted success.

Similarly, four of the seven subjects in the immediate condition who responded positively to the magic number seven question indicated they could remember 10 digits, while one of the five who had not heard about this predicted success. Eleven subjects claimed to have heard about research on detecting changes (although most of these were vague memories of TV shows or magazine articles describing continuity errors in movies). There was no difference in CBB between subjects who had (2.91 estimated change-detection successes) and had not (2.81 estimated change-detection successes) heard about change detection research.

In Experiment 2, knowledge about the digit span test had no effect on estimates (29% of 34 subjects who reported familiarity with digit span predicted success while 31% of subjects who were unfamiliar with the test did so). Knowledge about the magic number seven reduced estimates nonsignificantly (24% of the 21 subjects who knew about this predicted success while 33% of subjects who did not know about this effect predicted success, $\chi^2 = .56$).
Confidence ratings for positive estimates were generally high \((M = 5.45)\), and did not vary across conditions \(F < 1\).

**Attention question.** The total number of estimated change-detection successes was conditionalized on responses to the attention question. Overall, 20 subjects indicated that attention to the changing object was not necessary to detect a change, 16 indicated that attention was necessary, and 15 gave indeterminate responses such as “it depends on if the change is drastic or not.” The number of each subject’s predicted change-detection successes was entered into a mixed factors Attention Response (attention necessary/attention not necessary) \(\times\) Type of Change (identity/object) ANOVA to test for an interaction between the attention response and predicted success on different types of change. Neither the main effect for attention response nor the attention response \(\times\) type of change interaction was significant, \(F^{'s} < 1\).

**Digit span scenario.** In the immediate condition, 38% of subjects thought they could remember 10 digits whereas none of the subjects in either the minute or hour condition believed they could remember the 10 digits (minute-immediate comparison, \(p = .0064\); hour-immediate comparison, \(p = .0076\), Fisher’s exact tests).

### 2.3. Discussion

Overall, subjects vastly overestimated their change-detection ability in all conditions and were minimally affected by the memory reminder and the delay. Even when the delay was described as lasting one hour, estimates of success were only slightly less, and remained far above empirical baselines. Estimates in the immediate digit span scenario were also overoptimistic, but less so (for confirmation of this difference see the combined analysis in Experiment 2). However, in this scenario, subjects were strongly affected by the delay. The overestimate completely disappeared in the delay conditions where 0 of 38 subjects thought their memory could outlast a delay. At a minimum, this finding confirms that subjects were attending to the details of the scenarios. The attention results were also interesting in that responses were mixed regarding the necessity of attending to an object in order to notice a change, although these responses were unrelated to change-detection estimates.

Although the digit span estimates do suggest that subjects understood the delays in a general way, it is still possible that they were not salient enough to counter the

---

**Table 1**

Percentage of subjects predicting successful change detection in Experiment 1

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Immediate ((n = 13))</th>
<th>Minute ((n = 20))</th>
<th>Hour ((n = 18))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate</td>
<td>61.5%</td>
<td>35%</td>
<td>27%</td>
</tr>
<tr>
<td>Scarf</td>
<td>76.9%</td>
<td>75%</td>
<td>66.7%</td>
</tr>
<tr>
<td>Actor</td>
<td>61.5%</td>
<td>85%</td>
<td>66.7%</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>100%</td>
<td>95%</td>
<td>94.4%</td>
</tr>
<tr>
<td>Digit</td>
<td>38%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Means for change scenarios</strong></td>
<td><strong>75%</strong></td>
<td><strong>72.3%</strong></td>
<td><strong>64.8%</strong></td>
</tr>
</tbody>
</table>
easy visibility of the changes. In addition, subjects could have misconstrued the situations as presented in the scenarios and accompanying illustrations. For example, they may not have appreciated the transitory nature of the pre- and postchange views. Many of these interpretive issues arise from the fact that subjects in Experiment 1 made inferences about dynamic stimuli based on static images and verbal descriptions. Although this represents a good match to many real-world situations where metacognitive estimates are important, a closer match between the actual situation and its description would help eliminate potential alternative explanations for CBB. Therefore in Experiment 2, subjects estimated their change-detection ability based on videotapes showing the actual stimuli from the movie scenarios, and a subjects’ eye view of the pedestrian-change scenario.

3. Experiment 2

For Experiment 2, videos used in Levin and Simons (1997) illustrated each of the changes. For the pedestrian scenario, a subject’s-eye view of the switch with the original experimenters from Simons and Levin (1998) was used. A videotape of the digit span scenario was also included. In addition to the immediate condition, a 1-min delay condition was created in which the prechange and postchange views were separated by a shot of an actor experiencing the disruption. In the delay condition, subjects read the scenarios, viewed the prechange shot while the prechange object was pointed out, then saw an actor experiencing the disruption, and finally saw the postchange shot while the postchange object was pointed out. If CBB requires a large inference from static to dynamic stimuli, then it should disappear or be significantly reduced under these circumstances.

In addition, the postexperiment questionnaire was expanded to include questions asking subjects to justify their responses to each of the four change scenarios and the digit span scenario.

3.1. Method

Subjects. Seventy-three Kent State University undergraduates completed this experiment (37 in the immediate condition and 36 in the delay condition) in exchange for extra credit in their general psychology course. These subjects were run in two cohorts with two minor methodological changes described below.

Materials. In place of the illustrations used in Experiment 1, subjects were shown short videos from Levin and Simons (1997) and Simons and Levin (1998). For the plate and scarf scenarios, the pre- and postchange views were pairs of shots from Levin and Simons’ (1997) original 6-shot video. The actor-change video was originally only two shots, so it was played in its entirety. A video was created for the digit span task in which an actor read the 10 digits (1 per second) aloud. The delay videos were created by filming a new set of actors modeling the process of answering a phone while watching a movie (or hearing the 10 digits), or talking to a friend after having been interrupted while giving directions. The delay videos first showed the prechange shot. For the plate, scarf, actor, and digit scenarios, just
before the shot ended, a telephone ring was dubbed onto the soundtrack. After the phone finished its first ring, the first shot ended, followed by a brief blank screen (4–5 frames). Then, an actor watching the TV (or listening to the digits) could be seen stopping the VCR, walking away from the TV (or the testing table in the digit span scenario), answering the phone, and having a brief conversation. About 10s before 1 min had elapsed, the actor finished the conversation and returned to the TV or testing table. After this interruption there was another blank screen, followed by the postchange shot. In the digit scenario, the first shot showed an experimenter reading the 10 digits which was followed by a 1-min phone conversation.

There was one substantive difference between the digit span scenario and the others. In the digit span scenario, the 10 digits subjects were judging were visible on the screen while the experimenter read them, and remained visible during the delay. We did this to reduce the likelihood that any correct estimates of digit span would not be caused by actually forgetting the digits during the delay.

After the five scenarios, the first cohort of subjects’ digit spans were tested using the digit span subtask from the WAIS. In this task, an experimenter reads a series of digit strings of increasing length to subjects. After each digit string they were asked to write the numbers down in the correct order. Because this test only goes up to 9 digits, the second cohort of subjects were simply asked to remember a single 10-digit number.

The postexperiment questionnaire included all of the questions asked in Experiment 2 in addition to a set of five open-ended questions (one for each scenario) asking subjects why they responded the way they did for each scenario.

**Procedure.** Subjects were run in small groups ranging in size from 1 to 5. Subjects first read the same cover page used in Experiment 1, then completed the five scenarios by reading the text describing each and then viewing the video as an illustration. The first cohort of subjects viewed the scenarios in the following order: plate, scarf, actor, pedestrian, digit. The second cohort of subjects viewed the stimuli in the reverse order, starting with the digit scenario. While subjects viewed the videos, the experimenter pointed out the pre- and postchange objects with a laser pointer, saying “Here is/are x, and then here they change to x.”

### 3.2. Results

**Change detection scenarios.** As in Experiment 1, subjects showed CBB and again were minimally affected by the presence of a delay between pre- and postchange views. In the immediate condition they believed they would notice 70.8% of changes, and in the delay condition they thought they would notice 68.5% of the changes, \( F < 1 \). This represents a nonsignificant drop from the estimates in Experiment 1 in which subjects believed they would notice 75% of changes in the immediate condition and 72.3% in the 1-min delay condition, \( F < 1 \). Testing individually, the delay did significantly reduce estimates for the scarf scenario (78% immediate, 56% delay; \( X^2 = 4.307, p < .05; \) see Table 2) while it significantly increased estimates in the actor scenario (60% immediate, 83% delay; \( X^2 = 5.075, p < .025 \)). Delay had no effect on the plate and pedestrian scenarios. Again, in all scenarios in both
conditions, estimates were significantly greater than base rates (Fisher's exact test; \( p < .0001 \)).

Across all scenarios and both conditions the mean confidence rating for positive responses was 5.62 (compared with 5.45 in Experiment 1). Confidence ratings were 5.48 in the delay condition and 5.75 in the immediate condition, \( F(1,71) = 1.867, \ MSE = .693, \ p = .1761 \). The difference between immediate and delay conditions was driven by the pedestrian scenario for which subjects were significantly less confident in the delay condition (mean rating of 6.02) than the immediate condition (mean rating of 6.02) than the immediate condition (mean rating of 6.02) than the immediate condition (mean rating of 6.02) than the immediate condition (mean rating of 6.02) than the immediate condition (mean rating of 6.02).

**Attention question.** In Experiment 2, 21 subjects indicated that attention to the changing object was not necessary to detect a change, 23 indicated that attention was necessary, and 29 gave indeterminate responses such as “it depends on if the change is drastic or not.” Subjects who indicated attention was not necessary to detect changes predicted they would notice more changes (78.5%) than subjects who thought attention was necessary (65%; \( F(1,42) = 4.378, \ MSE = .358, \ p = .0478 \)). It appears that responses to the object changes underlie the attention response effect. Subjects who indicated that attention was not necessary thought they would detect significantly more object changes (69%) than subjects who thought attention was necessary (47.5%; \( F[1,42] = 4.172, \ MSE = .474, \ p = .047 \)). This effect was nonsignificant for the identity changes (88 and 82.5% predicted changes, respectively; \( F < 1 \)).

**Digit span scenario.** Subjects again overestimated their performance in the immediate digit span scenario (40.5% predicted success), and significantly reduced their estimates for the delayed scenario (19.4% predicted success; \( X^2 = 3.85, \ p < .05 \)).

When digit spans were actually measured (with no delay) using the Wechsler subtest, the mean span of the 37 subjects in the first cohort was 7.14 (\( SD = 1.51 \)), suggesting that 2.9% of subjects have a digit span of 10. None of the 36 subjects in the second cohort were able to remember a 10-digit number.

**Response justifications.** The open-ended response justifications were coded by two judges who rated the presence of six nonexclusive categories of responses for each question. First, the judges searched the responses for any mention of the need for memory in the task. Responses were also coded for mention of attention to (or memory for) (A) the specific changing object or (B) the category the changing object

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentage of subjects predicting successful change detection in Experiment 2</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Plate</td>
</tr>
<tr>
<td>Scarf</td>
</tr>
<tr>
<td>Actor</td>
</tr>
<tr>
<td>Pedestrian</td>
</tr>
<tr>
<td>Digit</td>
</tr>
<tr>
<td>Means for change scenarios</td>
</tr>
</tbody>
</table>

belonged to. For example, responses such as “I would have attended to/noticed the plates” and “I generally pay attention to/notice clothing” would fall into categories A and B, respectively. The fourth and fifth categories included responses indicating that the change would be noticeable (or not) due to its centrality within the scene, and referring to the subjects’ own capabilities or lack thereof. The final category includes references to the similarity of the pre- and postchange objects. Judges agreed on 91.8% of all ratings. Disagreements were resolved by discussion to produce a final set of ratings for analysis.

Open-ended responses revealed several interesting patterns. First, few subjects mentioned memory for the four change-detection scenarios while most mentioned it for the digit scenario. This holds true both for the immediate (78% mentioned memory in the digit scenario while 2.7, 2.7, 0, and 5.4% mentioned it in the plate, scarf, actor, and pedestrian scenarios, respectively, \(X^2's > 40.46, p < .001\) and delayed scenarios (81% mentioned memory in the digit scenario while 5.6, 2.8, 17, and 19% mentioned it in the plate, scarf, actor, and pedestrian scenarios, respectively, \(X^2's > 25.35, p's < .001\)). In addition, more subjects mentioned their own capabilities (or lack of capability) with reference to the digit scenario than all of the immediate change scenarios (43% in the digit scenario compared with 0, 11, 2.7, and 5.4% in the plate, scarf, actor, and pedestrian scenarios, respectively, \(X^2's > 9.87, p's < .005\), and for three of the four comparisons in the delayed condition (27% in the digit scenario compared with 5.5, 8.3, 2.7, and 22% in the plate, scarf, actor, and pedestrian scenarios, respectively, \(X^2's > 4.59, p's < .05\) for the plate, scarf, and actor scenarios; \(X^2 = .29\) for the pedestrian scenario).

Comparing among the change scenarios, it appears that similarity of pre- and postchange objects was relatively important for the identity-change scenarios (e.g., the actor and pedestrian scenarios), while general centrality and attention to the changing object were more important for object scenarios (see Table 3). In addition, fewer subjects mentioned similarity in the delayed identity change scenarios compared with the immediate identity change scenarios (actor scenario: 68% in immediate and 36% in delayed condition).

### Table 3

Percentage of subjects giving each category of response justifications for object-change and identity-change scenarios

<table>
<thead>
<tr>
<th>Response category</th>
<th>Scenario type</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Object change</td>
<td>Identity change</td>
</tr>
<tr>
<td>General centrality</td>
<td>63%</td>
<td>26%</td>
</tr>
<tr>
<td>Att/mem specific</td>
<td>52%</td>
<td>22%</td>
</tr>
<tr>
<td>Att/mem category</td>
<td>22%</td>
<td>36%</td>
</tr>
<tr>
<td>Similarity pre-post</td>
<td>12%</td>
<td>60%</td>
</tr>
<tr>
<td>Memory</td>
<td>4%</td>
<td>15%</td>
</tr>
<tr>
<td>Unique capability</td>
<td>11%</td>
<td>15%</td>
</tr>
</tbody>
</table>

*Note.* The percentages reflect the proportion of subjects who gave each response justification on the plate or scarf scenarios for object changes, and the actor or person scenarios. Between-condition differences were tested using \(\chi^2\).
delay, $X^2 = 7.23, p < .01$; pedestrian scenario: 54% in immediate scenario and 28% in delay scenario, $X^2 = 5.20, p < .025$).

Change detection responses for the actor scenario were conditionalized on similarity responses for the actor-scenario question to help understand why subjects thought they would be more successful in detecting changes across a delay in this case. There was no effect of delay on CBB for this scenario among subjects who indicated that similarity of the pre- and postchange actors was important (56% predicted noticing in the immediate scenario while 61% predicted success in the delayed scenario). Among subjects who did not mention similarity, 67% predicted success in the immediate scenario, while 96% predicted success in the delay scenario ($p = .0336$, Fisher’s exact test). Therefore it appears that subjects who did not consider the similarity of the actors in the delay scenario almost universally predicted successful change detection.

**Combined analysis of all experiments.** To get a better sense of the effects of delay for each scenario, data for immediate and delayed conditions were collapsed across Experiments 1 and 2 along with a third very similar experiment that does not appear in this report$^2$ (Fig. 1). Across all three experiments, the plate and scarf scenarios were associated with significantly less optimistic estimates in the delay conditions than in the immediate conditions ($X^2 = 6.033, p < .025$; $X^2 = 5.547, p < .025$, respectively), while the actor scenario was associated with a nonsignificant trend for more optimistic estimates ($X^2 = 3.238, p < .10$), and the pedestrian scenario was associated with no change.

A comparison between responses to the digit scenario and the three change scenarios with similar base rates (the plate, scarf, and actor scenarios) confirms that overestimates were stronger for the change scenarios. (Comparisons for immediate scenarios: plate-digit, $X^2 = 5.67, p < .025$; scarf-digit, $X^2 = 32, p < .001$, actor-digit $X^2 = 14.29, p < .001$. Comparisons for delayed scenarios: plate-digit, $X^2 = 37, p < .001$; scarf-digit, $X^2 = 42.67, p < .001$; actor-digit; $X^2 = 57.2, p < .001$).

Finally, of the 203 subjects in these experiments, none predicted 0 successes, 8 (3.9%) predicted one success, 60 (29.6%) predicted two successes, 82 (40.4%) predicted three successes, and 53 (26.1%) predicted four successes. Thus, arguably 96.1% of our subjects showed CBB if we assume that one change-detection success is a reasonable estimate.

### 3.3. Discussion

Experiment 2 confirms a large CBB effect using dynamic stimuli that were very similar to those used to demonstrate change blindness originally. In addition,

---

$^2$ The additional experiment was similar to Experiment 1, and included a minute-delay condition ($n = 28$), and two immediate conditions, one in which subjects were told about the importance of memory in combining views ($n = 23$), and one in which they were not told this ($n = 28$). Subjects predicted 73.3% success in the delay condition, 80.25% success in the immediate/memory condition, and 72.8% success in the immediate/no memory condition, $F(2,76) = 1.012$, ns.
Experiment 2 again shows no overall effect of delay on estimates although subjects’ confidence ratings were significantly lower in the delay condition. Experiment 2 also suggests that explicit beliefs about visual memory are not central in explaining CBB. In contrast to the majority of subjects who mentioned memory in the digit span scenario, few mentioned it for the CBB scenarios. Beliefs about the necessity of attention to detect changes seem more relevant although they were only loosely related to CBB. In both Experiments 1 and 2, subjects were split on the necessity of attending to the changing object in order to see changes, and showed CBB irrespective of their response to the attention question. However, in Experiment 2, subjects who thought attention was unnecessary to detect changes did show stronger CBB.

In addition to beliefs about attention, variables related to the structure of the scene and the perceptual similarity of the pre- and postchange objects appear to underlie CBB. For the object changes, the most frequently given response justifications emphasized the centrality of the object in the scene and the degree to which they would pay attention to the changing object. However for identity changes subjects indicated that similarity of the pre- and postchange objects was important. In a sense, this is unsurprising because identity changes are maximally central, almost by default. The finding does imply that subjects consider the relative status of objects in the scene, in some cases allowing their estimates to reflect the fact that
changes to noncentral objects will not necessarily be noticed. Even though this modification occurs relative to a large degree of optimism for both central and noncentral objects, it suggests that subjects use the structure of natural scenes as a guide for determining what changes they will notice.

Additional analyses showed that for the actor-change scenario, the paradoxical increase in optimism in the delay condition is a result of the combined importance of similarity and perhaps beliefs about identity tracking. This effect was limited to the subset of subjects who indicated that similarity was not important in their decision. Accordingly, it is likely that subjects strongly believe that they will detect the actor substitution, even after a delay, unless they notice that the two actors are quite similar. This pattern of results suggests that in the absence of information to the contrary subjects default to a belief that violations of continuing identity, even over the course of a minute or so, are automatically detected.

The overall analysis of all three experiments reinforces the prevalence of CBB, and in addition confirms that it is stronger than overestimates of digit memory. It is important to note that we included the digit span scenario as a rough index of a general cognitive overestimate, and not for a precise comparison between CBB and metamemory which we will leave to future research. For now, we note that we previously observed stronger overestimate digit performance when subjects predicted their own performance than when they predicted that of others (estimate of self-success: 40%, estimate of other-success: 19%), perhaps due to the desire for positive self-presentation, while CBB was not sensitive to this manipulation (Levin et al., 2000). This is reinforced by the present experiment in which subjects were far more likely to mention their skills and abilities with reference to the digit scenario than the change scenarios. Therefore, the self-estimates tested here afford a particularly large overestimate for the digit scenario, but not the for CBB scenarios. Despite this, CBB was still stronger.

4. General discussion

These experiments show that change blindness is robust and can be observed even when it is unlikely that a perceptual transient would signal the change. Both memory reminders and the insertion of delays of up to an hour had only minimal overall impact on the effect. In addition, subjects showed strong CBB even when they viewed videos depicting the changes. When effects of delay are broken down by scenario, it becomes clear that delays reduced estimates for the plate and scarf scenarios, did not reduce estimates for the pedestrian-change scenario, and in some cases actually increased estimates for the actor scenario. Analyses of responses to open-ended questions suggest that some subjects believe that attention to the changing object is necessary to detect a change while some think that the change will pop out at them. However, both groups of subjects show CBB. Finally, CBB was in all scenarios stronger than overestimates of digit span.

Given these findings it is unlikely that CBB is caused by a misconstrual of the perceptual experience associated with a change. As mentioned in the introduction, subjects may have believed that changes are associated with an easily perceptible
transient. However, the fact that they continue to make this error even when this transient would plainly not occur suggests that CBB reflects a more general belief about vision. This is important because it emphasizes the fact that CBB is not a situation where subjects have an essentially correct metacognitive understanding but make mistakes when they must project it into a situation that is unfamiliar, or in which subtle task-specific factors cause unpredictable breakdowns in performance. The scenarios used here are familiar on their surface: Everyone knows what it is like to watch a movie or talk with someone on the street. In addition, many people are familiar with the idea that changes might occur in movies and escape notice. Thus, subjects have good support for their predictions, and should therefore have every chance to make correct estimates. The fact that they do not suggests that their understanding of vision conflicts deeply with visual functioning.

Combined, these findings not only reinforce the range of situations over which CBB occurs, but also have implications for explaining the effect. One thing these experiments suggest is that different beliefs contribute to CBB for identity changes and CBB for object changes. As discussed in the introduction, this distinction is related to previous distinctions between change detection in central and noncentral objects (Levin & Simons, 1997; Rensink et al., 1997).

A variety of developmental findings suggest both a perceptual and conceptual tendency to maintain the continuity of centrally attended objects over time, even in the face of conflicting property information. (see, for example, Kellman & Spelke, 1983; Xu & Carey, 1996). Particularly relevant here is Gutheil and Rosengren’s (1996) argument that by age 4, children judge that objects retain their identity over transformations, suggesting that they conceptualize a continued history of time and place as a sign of identity independent of specific visual features. It is particularly interesting to note that in this example, children inferred continued identity by assuming that a distinctive behavior produced by an animal would continue even if the animal’s appearance changed. In a sense this is similar to the identity changes we studied here because the observer’s putative goal is to follow a behavioral narrative as different views of people are combined to produce one continuous event. Accordingly, behavioral consistency constitutes the thread that combines views, and if subjects fail to realize this they may falsely infer that visual property information is the basis of continuity. Centrally to this claim is the idea that subjects might believe they track visual details specifying identity per se, and not just any visual detail characteristic of an attended object or person. For example, it is possible that subjects do not think they track clothing over time as part of their identity tracking process—people should be aware that they might not notice or care about someone wearing a different shirt from one day to the next.

Although subjects clearly overestimate their ability to detect changes in attended objects, they also overestimate their ability to detect changes to less central objects that are not the focus of attention. In fact, the magnitude of this overestimate is, in most cases, larger than that for identity changes because the high estimates are paired with empirical success rates of essentially 0%. However, CBB for object changes is distinct from that for identity changes in a number of ways. First, object-change CBB is reduced slightly, but significantly by the pre- to postchange delays. It is important to note that although delay causes some reduction in CBB, this is not
due to a straightforward application of explicit beliefs about the necessity of memory for change detection. In the postexperiment response justifications, very few subjects mentioned memory for the object scenarios, while most did for the digit span. Instead, the response justifications suggest that subjects focus on the relationship between the changing object and the rest of the scene indicating that they would notice central items and fail to notice noncentral items. Although we cannot be certain about why the delay reduced object estimates, it is possible that the delays disrupted what was otherwise a coherent event. If subjects analyze visual organization at the level of events in these videos, then it is possible that, as discussed below, CBB was reduced because visual organization was effectively reduced.

What, then, underlies CBB for object changes? Two specific factors seem like good candidates at this point. First is a misconstrual of the typical course of attention in scene perception. At least some subjects appear to believe that they typically attend to all of the salient objects in a scene, and can therefore detect changes in the scene. Many subjects indicated that they would have to pay attention to changing objects to see them change, while also claiming they would, in fact, see at least one of the object changes. For example, in Experiment 2, 74% of subjects who indicated that attention was usually necessary to see a change also indicated that they would see at least one of the object changes (35% of these subjects claimed they would see both). This suggests that subjects believe they somehow consciously process most objects in the scene (Levin, 2002). The tendency to respond as if this were true could be enhanced by a hindsight bias in which people have difficulty simulating ignorance about the location of the change (Fischhoff, 1982).

In addition to misunderstanding of the typical course of attention and hindsight, the well-organized nature of our scenes may lead subjects to believe either that they can represent the entire scene, or that the scene is essentially a single memorable “chunk” that can be represented as a whole. A number of response justifications give this sense. For example, one subject wrote, “I would often notice the change without paying direct...attention to the object but the entire scene itself,” and another wrote, “I think you could see the changes without paying direct attention to the objects because the changes often affected the whole scene.” Thus subjects may assume that the inter-object relationships characteristic of natural scenes allows them to efficiently code and compare details across views. This assumption may be related to the understanding that elaboration and organization are useful memory strategies (see, for example, Justice & Weaver-McDougall, 1989; Sodian & Schneider, 1986), and that it is easier to focus attention, for purposes of memorization, on a subset of pictures if that subset forms a category (Miller & Weiss, 1982). It is interesting to note that change detection may, in some cases, not be facilitated by this kind of organization. For example, Mitzumatsu and Yokosawa (2000) report that jumbling natural scenes does not always interfere with change detection.

4.1. Why study visual metacognition?

Before concluding, we would like to briefly discuss the importance of understanding the kinds of visual metaknowledge we have explored here. In doing so, we
follow a basic distinction central to the metamemory literature: that between metaknowledge (e.g., people’s beliefs about or model of their own cognitive systems) and metacognitive control (e.g., executive control over cognitive processes; see for example Fernandez-Duque, Baird, & Posner, 2000). Accordingly, understanding beliefs about vision are important both for the potential they have to guide performance of visual task, and in themselves. First, similar to beliefs about memory, beliefs about visual function can affect visual performance, especially where visual attention is strategically allocated, or allocated based on concepts and/or knowledge that is less deliberatively accessed. If there is one thing that change blindness tells us, it is that much of visual experience is based on just this kind of selective knowledge-driven visual coding. One dramatic illustration of this issue was observed while testing heads-up displays intended to make airline piloting easier by projecting navigation-relevant information on the windshield of the cockpit so that pilots do not have to glance away from the outside world at important moments (Haines, 1991). Although pilots liked the system, they sometimes landed their plane right on top of another airplane taxiing onto their runway. Why did they do this? One reasonable explanation is that they do not understand visual attention—they do not realize that attending to the heads-up display precluded to noticing large objects in the same region of the visual field. Therefore they failed to search the scene for unexpected changes. It is important to note that pilots are probably well aware that looking down at their instruments does preclude them from seeing outside their windshields. So, it would seem worthwhile to specify where, exactly, this explanatory gap lies if we are to have a principled understanding of this kind of visual error.

In addition to the potential relationship between metaknowledge and visual performance, beliefs about visual metacognition are important to understand in themselves. One critical reason for this is that these beliefs guide people’s evaluation of others’ psychological experience. Recently, a number of compelling examples of this kind of evaluation have made the headlines with regard to metamemory. In a review of 40 convictions overturned based on DNA evidence, Wells et al. (1998) found that 90% of these errors were based on mistaken eyewitness identifications (5 of these individuals had been sentenced to death). Wells et al. reviewed extensive evidence showing that people falsely believe that there is a strong confidence-accuracy correlation, and argued that this false belief leads jurors to give far more weight to eyewitness testimony than is warranted. Thus, it appears that incorrect metacognitive beliefs can lead jurors to falsely convict people based on confident eyewitness testimony. In another applied context, computer programmers appear to believe that they understand users’ visual limits and can therefore determine a priori if an interface is usable. However, when actually tested, programmers can be at chance when distinguishing well designed interface displays from poorly designed interfaces (reviewed in Landauer, 1995). Therefore, if we can understand these individuals’ visual metacognitive beliefs, we can focus on educational strategies that point to specific situations where programmers intuitions are most likely run afool of their users capabilities.

Finally, it is important to clearly demonstrate that phenomena discovered in the behavioral sciences do, in fact, sometimes conflict with intuition. Recently, a colleague described an experience he had as an expert witness in visual perception. He
was not allowed to testify about change blindness because the judge claimed that all of the phenomena he was about to discuss “were consistent with intuition” and therefore did not require an expert. If there is one thing CBB tells us, it is that CB is not consistent with the intuitions of at least 96% of the subjects we tested in these experiments. In addition, although it is intuitive to suggest that CBB reflects a global belief in detailed representations, based on the apparent continuity of experience and easy availability of visual detail (Noë, Pessoa, & Thompson, 2000; O’Regan, 1992; Rensink, 2000), our data suggest that the phenomenon is much richer than that. CBB reflects a set of beliefs about the normal course of scene viewing and object tracking in natural scenes. These probably include beliefs about identity tracking, attention, and the degree to which the coherence of natural scenes allows efficient representation. Combined, these beliefs lead to a powerful and almost universal misconstrual of the degree to which changes in visual scenes can be detected.

Acknowledgments

We thank Bill Merriman and Julia Noland for reading and commenting on previous versions of the manuscript. Portions of this work were presented at the 1999 Vision Research conference in Ft. Lauderdale.

Appendix A. Sample scenarios from Experiments 1 and 2

A.1. Scarf scenario—immediate condition

Imagine that you are watching a movie. In a certain scene, two actors are having a conversation in a restaurant. In a series of shots you see one actor, then the other as each in turn talks. In one shot, actor A is wearing a large colorful scarf. In the next shot, actor A is shown from a different angle and the scarf is not present. You did not expect this change beforehand.

Would you notice the disappearance of the scarf?

A.2. Scarf scenario—delay conditions

Imagine that you are watching a movie. In a certain scene, two actors are having a conversation in a restaurant. In a series of shots you see one actor, then the other as each in turn talks. In one shot, actor A is wearing a large colorful scarf. While you are watching this shot, you stop the VCR and go to another room to answer the telephone. The call is from your friend who has a brief question for you, and you talk on the phone for 1 min. (Exp. 2, hour condition: You end up talking on the phone to your friend for one hour.) After hanging up the phone, you play the movie again and it starts on the next shot. In this shot, actor A is shown from a different angle and the scarf is not present. You did not expect this change beforehand.

Would you notice the disappearance of the scarf?
References


