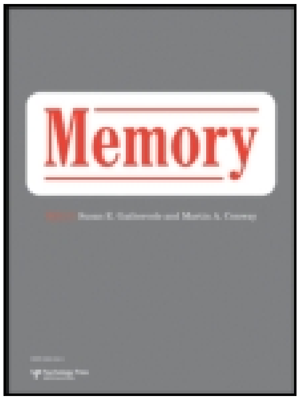

Memory
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pmem20

Change detection inflates confidence on a subsequent recognition task

Ryan J. Fitzgerald [a] , Chris Oriet [a] & Heather L. Price [a]
[a] Department of Psychology, University of Regina, Regina, SK, Canada

Available online: 21 Oct 2011

To cite this article: Ryan J. Fitzgerald, Chris Oriet & Heather L. Price (2011): Change detection inflates confidence on a subsequent recognition task, Memory, 19:8, 879-890

To link to this article: http://dx.doi.org/10.1080/09658211.2011.613843

MEMORY, 2011, 19 (8), 879–890

Psychology Press
Taylor & Francis Group

Change detection inflates confidence on a subsequent recognition task

Ryan J. Fitzgerald, Chris Oriet, and Heather L. Price

Department of Psychology, University of Regina, Regina, SK, Canada

A face viewed under good encoding conditions is more likely to be remembered than a face viewed under poor encoding conditions. In four experiments we investigated how encoding conditions affected confidence in recognising faces from line-ups. Participants performed a change detection task followed by a recognition task and then rated how confident they were in their recognition accuracy. In the first two experiments the same faces were repeated across trials. In the final two experiments novel faces were used on each trial. Target-present and target-absent line-ups were utilised. In each experiment participants had greater recognition confidence after change detection than after change blindness. The finding that change detection inflates confidence, even for inaccurate recognitions, indicates recognition certainty can be a product of perceived encoding conditions rather than authentic memory strength.

Keywords: Change detection; Change blindness; Memory; Confidence; Face recognition; Confidence–accuracy relation; Optimality hypothesis.

Change blindness is the failure to detect changes in a scene. Although this task has been studied extensively in the laboratory (e.g., Rensink, O'Regan, & Clark, 2000), much less attention has been paid to its possible role in other contexts. Davies and Hine (2007) among others have recently explored one such role, the effects of change detection on eyewitness memory for a crime. Their results and other findings in the well-established eyewitness memory literature provide a useful starting point for exploring the consequences of change detection failure more generally. Thus we begin our investigation by highlighting some of these findings before presenting the results of four laboratory experiments that suggest successful change detection inflates confidence in recognition memory.

In the context of eyewitness memory, change blindness can manifest as a failure to discriminate between two or more people who were observed at the scene of the crime, leading the witness to report the presence of a single person where there was in fact more than one. The difference between the memory of an eyewitness who experiences change blindness and an eyewitness who experiences change detection (i.e., successfully discriminating between two people observed at a crime scene) can be characterised as a

Address correspondence to: Ryan J. Fitzgerald, Department of Psychology, University of Regina, 3737 Wascana Parkway, Regina, SK, Canada, S4S 0A2. Tel: (306) 585-5339. E-mail: fitzgerr@uregina.ca

This research was supported by a Canadian Graduate Scholarship from the Social Science and Humanities Research Council of Canada to the first author and by grants from the National Sciences and Engineering Research Council of Canada to the second and third authors. Portions of this work were presented at the 71st Annual Convention of the Canadian Psychological Association, June, 2010, Winnipeg, Canada, and at the 20th Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, June, 2010, Halifax, Canada.

We would like to thank J. Don Read for the thoughtful insight and direction he provided us throughout the course of this project. In addition we express our gratitude to two anonymous reviewers for helpful comments made on a previous version of the manuscript.

http://www.psypress.com/memory http://dx.doi.org/10.1080/09658211.2011.613843

discrepancy in memory resulting, at least partially, from differences in the quality of encoding conditions. Lindsay, Read, and Sharma (1998) found that encoding conditions (in this case: exposure duration, perspective, and viewing instructions) were predictive of eyewitness confidence in line-up identifications. Therefore recognition confidence could be influenced by whether or not a change is detected. Here we report the results of four experiments that clearly show change detection increases confidence in a subsequent recognition task, and that this increased confidence is independent from improved performance in discriminating the target from foils.

The relation between line-up identification accuracy and confidence has been studied extensively. Bothwell, Deffenbacher, and Brigham (1987) conducted a meta-analysis of 35 eyewitness studies and found an aggregate confidence–accuracy correlation of .25. In another meta-analysis of 30 studies Sporer, Penrod, Read, and Cutler (1995) found a comparable confidence–accuracy correlation of .29. Although these meta-analyses indicate a somewhat modest relation between confidence and accuracy, there is evidence to suggest its strength can be influenced by encoding conditions.

Deffenbacher (1980) reviewed 43 eyewitness studies and, based on this analysis, formulated the optimality hypothesis to explain the relation between accuracy and confidence. According to Deffenbacher, the confidence–accuracy relation is moderated by information-processing conditions (i.e., how an event is encoded, stored, and retrieved). When conditions are poor, the confidence–accuracy relation should be weak. When conditions are good, the confidence–accuracy relation should be moderate to strong. Bothwell et al. (1987) found support for the optimality hypothesis in their meta-analysis, noting the length of exposure to a target accounted for 27% of the variability in predicting accuracy from confidence.

Target exposure duration has been examined in numerous studies since Bothwell et al. (1987) first identified it as an important moderator of the confidence–accuracy relation. In one such experiment (Read, 1995) research assistants posed as customers and interacted with store clerks for short or long durations. When asked to identify the customers from a line-up, clerks who had long interactions were more confident than clerks who had short interactions for both correct and incorrect identifications. Consequently, the confidence–accuracy relation was stronger for short interactions than for long interactions. In another experiment (Memon, Hope, & Bull, 2003) exposure duration was manipulated via staged crime videos. When participants had less exposure to the perpetrator they made correct identifications more confidently than false identifications and incorrect rejections. In contrast, when participants had more exposure to the perpetrator accuracy had no effect on confidence. Contrary to the optimality hypothesis, these experiments suggest better encoding conditions weakened the confidence–accuracy relation.

Lindsay et al. (1998) found partial support for the optimality hypothesis by manipulating video duration, viewing perspective, viewing instructions, and retention interval. Consistent with the optimality hypothesis, favourable information-processing conditions had a positive impact on accuracy and confidence, whereas less-favourable conditions had a negative impact on accuracy and confidence. However, the data for false identifications were inconsistent with the optimality hypothesis. Specifically, as encoding conditions improved, confidence increased for false identifications in the same manner as it did for correct identifications (see also Lindsay, Nilsen, & Read, 2000, for similar results). Lindsay et al. concluded that witnessing conditions influenced meta-memorial judgements about the likelihood of making a correct identification, leading participants to base their confidence on the quality of encoding conditions.

The experiments outlined above suggest encoding conditions do not always have parallel effects on confidence and accuracy. The optimality hypothesis seems to be valid in the case of correct identifications. Indeed, accurate identifications made under good encoding conditions tend to be made confidently. However, confidence also tends to be high for incorrect identification decisions when encoding conditions are good. In other words, regardless of identification accuracy, good encoding conditions seem to elicit a high degree of confidence.

In past research on the influence of encoding conditions on eyewitness memory, external viewing conditions were manipulated by the researchers. One drawback to this approach is that the fidelity of the memory for the perpetrator might be confounded with the extent to which details other than the perpetrator are attended. For example, as the length of a video increases, so do the opportunities to attend to details other than the perpetrator's identity. Better attention to the scene would yield the pattern observed by previous researchers, namely an increase in correct recognitions (resulting from better encoding of both the perpetrator and the scene) and an increase in false recognitions (resulting from better encoding of the scene only), coupled with increased confidence (resulting from the increased ease with which specific details come to mind and a general sense that the scene is remembered well). Alternatively, participants' motivation may have differed across poor and good encoding conditions. For example, participants who view a crime scene that is poorly lit or from a single perspective might not be as engaged by the task as participants who view the scene under better conditions, or might not be as motivated to be accurate, believing that their attempts to identify the perpetrator are unlikely to succeed.

In the present research the difference between encoding conditions was based on participants' actual performance (i.e., whether changes in stimuli were detected or not). In order for change detection to occur, sufficient encoding of the pre- and post-change stimuli is necessary. Furthermore, change detection is more likely to occur when participants are actively attending to the stimuli (Simons & Levin, 1997). Thus we assert that participants should perceive the viewing conditions as better if the change was detected than if the change was not detected. Importantly, this measure of the perceived quality of the encoding conditions is independent of recognition memory performance. As such, confidence in the recognition memory task can be studied as a function of the perceived quality of encoding conditions.

Change blindness has been studied using a variety of methods. In the flicker paradigm (e.g., Rensink et al., 2000), the method used in the present research, participants are shown a sequence of rapidly alternating displays. Typically, a single cycle consists of four brief displays: (1) a pre-change display; (2) a blank screen; (3) a post-change display; and (4) a blank screen. Another method, and one that is arguably more similar to change blindness in the "real world", uses the video paradigm. In one example, Levin and Simons (1997) showed a video that started with a man getting out of his chair to answer a telephone. In the following scene a different person was shown picking up the receiver. Change blindness has also been observed in live interactions. In one study (Levin, Simons, Angelone, & Chabris, 2002) participants who had been talking to a confederate often failed to notice when the confederate ducked underneath a desk and was replaced by a new confederate. Thus the phenomenon of change blindness is not limited to laboratory experiences.

In the present research we report the findings of four experiments. In each experiment participants performed a change detection task, followed by a recognition judgement, and then they rated how confident they were in their recognition judgement. Our primary interest was to evaluate the effect of different encoding conditions on recognition confidence. Although identification confidence was assessed in one video-based change detection study (Nelson et al., 2011), statistical comparisons between the change blindness and change detection groups were precluded by an exceptionally low rate of change detection (5%). Therefore it is presently unclear how awareness of change will influence recognition confidence. Based on previous studies in which confidence was higher for good encoding conditions than for poor encoding conditions (Lindsay et al., 1998, 2000; Read, 1995), we hypothesised that change detection (i.e., good encoding condition) would result in greater confidence than change blindness (i.e., poor encoding condition). In addition we anticipated that accurate recognition judgements, when possible, would be accompanied by greater confidence ratings than inaccurate recognition judgements. Lastly, in accordance with the optimality hypothesis, we predicted that the confidence–accuracy relation would be stronger when change was detected compared to when it was not.

EXPERIMENT 1

Method

Participants. A total of 26 university undergraduates participated in exchange for either partial course credit or payment of $10.

Materials. The stimulus set consisted of nine Caucasian male face images. The same faces were presented on each trial. Eight of the faces were photographs of people who were paid an honorarium for the use of their images (a $5 gift card). One of the faces was obtained from the Eberhardt

Lab Face Database. The images were in greyscale and cropped so that only the person's face was visible (i.e., no neck or shoulders).

Procedure. Participants completed 60 trials of a change detection task, followed by a recognition task. In the change detection task an array of six faces was displayed for 2 seconds, then a blank screen was shown for 200 ms, followed by a second array of six faces that was displayed for another 2 seconds. Five of the faces in the second display were the same as in the first display and one was different. Participants were instructed to identify the location of the face that changed. Feedback was provided to inform the participants of whether or not they correctly identified the location of the change. In the recognition task participants were shown three faces. One of the faces was present in one or both of the two previously viewed arrays of faces and two were not. The correct recognition choice consisted of either the face from the first display that changed into a new face in the second display (pre-change face), the new face from the second display (post-change face), or one of the other five faces that did not change (unchanged face). After the recognition judgement was made, participants were asked to rate their confidence on a scale of 1 (not at all confident) to 5 (very confident). Individual trials were separated by a blank screen that was shown for 3 seconds. No feedback was provided for the recognition judgement.

The three types of correct recognition choices (pre-change, post-change, or unchanged faces) were randomised across trials. The positioning of the stimuli, the location of the face that changed, and the positioning of the correct recognition choice were also counterbalanced across trials. This was achieved by creating a trial for every possible combination of these three variables and randomly selecting from the resulting set of trials. For trials on which the correct recognition choice was a pre-change face, participants completed 20 trials that were randomly selected from a list of 162 trials (9 stimuli arrangements $\times$ 6 locations of the face that changed $\times$ 3 locations of the correct recognition choice). Similarly, for trials on which the correct recognition choice was a post-change face, participants completed 20 trials that were randomly selected from a list of 162 trials. For trials on which an unchanged face was the correct recognition choice, participants completed 20 trials that were randomly selected from a list of 810 trials (5 unchanged faces $\times$ 9 stimuli arrangements $\times$ 6 locations of face that changed $\times$ 3 correct recognition choice arrangements).

Results

The location of the face that changed was correctly identified on 54.2% of all trials, yielding a comparable number of trials on which change detection and change blindness occurred. Familiarity effects were assessed by dividing the 60 trials into four quartiles and comparing recognition accuracy and confidence among the four quartiles. A repeated-measures ANOVA revealed trial quartile had no effect on recognition accuracy, $F(3,\ 75) < 1$, or on confidence, $F(3, 75) = 1.14$, $p = .34$, $\eta_p^2 = 04$.

A 2 (recognition accuracy: incorrect vs correct) $\times$ 2 (awareness: change blindness vs change detection) repeated-measures ANOVA was conducted with confidence as the dependent variable. Confidence was higher after change detection than after change blindness (see Table 1), yielding a significant main effect of awareness, $F(1,\ 25) = 18.03$, $p < .001$, $\eta_p^2 = .42$. Confidence was also higher following correct recognitions (i.e., indicating that a face that was present in the previous displays was present in the line-up,

TABLE 1
Mean confidence ratings as a function of awareness of change and recognition accuracy

	Change blindness			*Change detection*		
Exp.	*Incorrect recognition*	*Correct recognition*	*Total*	*Incorrect recognition*	*Correct recognition*	*Total*
1	2.94 (0.72)	3.26 (0.77)	3.03 (0.68)	3.08 (0.73)	3.65 (0.74)	3.40 (0.70)
2	2.70 (0.91)	–	2.70 (0.91)	3.03 (0.86)	–	3.03 (0.86)
3	2.33 (0.59)	2.99 (0.86)	2.61 (0.64)	2.74 (0.72)	3.56 (0.62)	3.19 (0.63)
4	2.81 (0.59)	3.18 (0.71)	2.95 (0.57)	3.18 (0.65)	3.70 (0.70)	3.39 (0.63)

Standard deviations are in parentheses.

$M = 3.50$, $SD = 0.71$) than following incorrect recognitions (i.e., indicating that a face that was absent from the previous displays was present in the line-up, $M = 3.00$, $SD = 0.69$), leading to a significant main effect of accuracy, $F(1, 25) = 46.24$, $p < .001$, $\eta_p^2 = .65$. The interaction was not significant, $F(1, 25) = 2.31$, $p = .14$, $\eta_p^2 = .08$. Paired samples *t*-tests showed confidence was higher for correct recognitions than for incorrect recognitions both on change detection trials, $t(25) = -6.52$, $p < .001$, $d = 1.28$, and on change blindness trials, $t(25) = -2.62$, $p = .01$, $d = 0.52$. Thus the effect of accuracy on confidence was observed irrespective of whether or not the change was detected.

Discussion

Confidence was higher for change detection trials than for change blindness trials. Thus our first hypothesis was supported. This finding was interpreted as evidence that feedback about the viewing conditions influenced participants' certainty for the subsequent recognition decision. The inflation of confidence when encoding conditions were good compared to poor fits well with previous studies that have examined the role that encoding conditions play in witness confidence (Lindsay et al., 1998, 2000; Memon et al., 2003; Read, 1995). In contrast to these studies, however, the present work used a within-participants design, precluding the possibility that the manipulation of encoding conditions itself influenced participants' level of engagement in the task, or the details to which they attended. Consistent with the conclusions reached by previous authors, the participants' experience of the encoding conditions as poor or good appeared to influence their confidence in the subsequent recognition memory judgement.

In support of the second hypothesis, participants were more confident when they made correct recognitions than when they made incorrect recognitions. The third hypothesis, namely that the confidence–accuracy relation would be stronger when encoding conditions were good compared to when encoding conditions were poor, was also supported. Although confidence was higher for correct recognitions than for incorrect recognitions on both change detection and change blindness trials, the effect sizes differed substantially. According to the conventions set by Cohen (1988), the strength of association was large when the change was detected ($d = 1.28$) and only medium when the change went undetected ($d = 0.52$).

EXPERIMENT 2

In the first experiment a correct recognition choice could always be made. The use of a target-present recognition task was important for determining the confidence–accuracy relation. One limitation of the target-present-always design, however, is that participants may have experienced a sense of familiarity at the time of viewing the choices for the recognition memory task, which in turn may have artificially increased their confidence in their choice. Therefore in the second experiment participants were given a target-absent recognition task. By using only target-absent trials, we were able to test an intriguing hypothesis: Would successful change detection increase confidence in a recognition decision that could not possibly be correct? Such a finding would suggest that successful change detection leads to an inflated impression of the quality of encoding conditions, without actually increasing the likelihood that the changed target can be successfully discriminated from distractors. In addition, if change detection increases confidence on a task on which no correct answer can be given, it will show that the effect of encoding conditions on confidence occurs independently from recognition accuracy.

Method

Participants. A total of 26 university students were recruited from undergraduate psychology courses. They received partial course credit for their participation. None had participated in any other experiment reported in this paper.

Materials. The stimulus set consisted of the nine face images from Experiment 1 and one new face image obtained from the Eberhardt Lab Face Database. All faces were of Caucasian males.

Procedure. Participants completed 90 trials that were identical to Experiment 1 with the exception that the recognition task was target-absent. That is, participants were shown three faces and asked to identify the one that was in either of the two previous displays, but in reality none of the three faces was shown in either of the previous two displays. The interstimulus interval (ISI) was manipulated between the first and second displays, and was chosen randomly on each trial to be 100, 200, or 300 ms. Using the same procedure as in Experiment 1, the positioning of the stimuli, the location of the face that changed, and the positioning of the recognition choices were counterbalanced across trials.

Results

The location of the change was detected on approximately half of all trials (50.3%), ensuring a comparable number of change detection and change blindness trials for analysis. A repeated-measures ANOVA showed ISI had no reliable effect on change detection accuracy, $F(2, 50) = 2.60$, $p = .08$, $\eta_p^2 = .09$. A 2 (awareness: change blindness vs change detection) $\times$ 3 (ISI: 100 ms vs 200 ms vs 300 ms) repeated-measures ANOVA with confidence as the dependent variable revealed a main effect of awareness, $F(1, 25) = 23.22$, $p < .001$, $\eta_p^2 = .48$, with no main effect of ISI ($F < 1$) and no interaction ($F < 1$). As in Experiment 1, confidence was higher after change detection than after change blindness (see Table 1).

Discussion

The results of Experiment 2 demonstrate that perceiving the encoding conditions as good inflated confidence in a task in which confidence should be very low because no correct choice was available. Because no correct answer could be made, this finding indicates the inflation of confidence on change detection trials versus change blindness trials was independent of recognition accuracy.

EXPERIMENT 3

In the first two experiments feedback on change detection performance was provided and the same faces were displayed repeatedly across trials. Arguably, these methodological artefacts might have confounded the results. Although we interpreted the results as evidence that change detection led to higher confidence than change blindness, an alternative interpretation is that seeing the word ‘‘correct’’ after the change detection task led to higher confidence than seeing the word ‘‘incorrect’’ after the change detection task. Because participants received feedback after the change detection task on every trial, it is unclear whether it was the detection of change or the positive feedback that caused the increase in confidence. The results could also have been influenced by the faces becoming more familiar as the experiment progressed. Although we found no evidence of improved recognition accuracy as Experiment 1 progressed, familiar faces are undoubtedly processed and represented differently than unfamiliar faces (Megreya & Burton, 2006, 2007). For example, although participants are poor at verifying the identity of unfamiliar faces seen on video, they do well at this task when the faces are familiar (Bruce, Henderson, Newman, & Burton, 2001). One problem associated with the use of repeated faces is that participants may then be able to develop a face-naming strategy (e.g., label one face as ‘‘the thin guy’’).

Experiment 3 was designed to investigate whether or not confidence would be affected by awareness of change when a face-naming strategy would not be likely to be useful. This was accomplished by using new faces for each trial. In addition, no feedback was given on the change detection task. Therefore, if confidence is higher after change detection than after change blindness, we can be sure that the act of detecting the change itself is causing the inflation in confidence, as opposed to the feedback provided in the first two experiments. One additional change in Experiment 3 is that a mix of target-present and target-absent trials was employed. This design allowed us to examine line-up type as a within-participant factor.

Method

Participants. A total of 21 university undergraduates were recruited from psychology courses. They received partial course credit for their participation.

Materials. A total of 333 Caucasian male faces were used as stimuli. The faces were obtained from a collection of face databases, including the Eberhardt Lab Face Database, the Center for Vital Longevity Database (Minear & Park, 2004), and the Facial Recognition Technology (FERET) Database. Some faces from the previous two experiments were used as well. As in the first two experiments all faces were cropped to allow only the face to be seen. In contrast to the first two experiments, colour images were used.

Procedure. Participants completed 36 trials (27 target-present; 9 target-absent) that were similar to the first two experiments (a change detection task, followed by a recognition task, and then a confidence judgement). For the change detection task the first and second displays were separated by a 200-ms inter-stimulus interval and no feedback was provided. For target-present trials the correct recognition choice randomly varied among pre-change faces, post-change faces, and unchanged faces (nine trials of each). An effort was made to control for naming strategies within trials by ensuring that the faces shown in the change detection task were at least superficially similar to each other, and similar to the faces in the recognition task (e.g., all with beards or all clean-shaven).

Results

The location of the face that changed was correctly identified on 56.6% of all trials. A 2 (line-up type: target-present vs target-absent) $\times$ 2 (awareness: change blindness vs change detection) repeated-measures ANOVA was performed on confidence ratings. This test revealed higher confidence for change detection trials than for change blindness trials (see Table 1), yielding a significant main effect of awareness, $F(1, 20) = 11.32$, $p = .003$, $\eta_p^2 = .36$. Paired samples *t*-tests confirmed confidence was higher after change detection than after change blindness for both target-present trials, $t(20) = -3.50$, $p = .002$, $d = 0.88$, and for target-absent trials, $t(20) = -2.13$, $p = .045$, $d = 0.50$. In addition, confidence was higher on target-present line-up judgements ($M = 3.13$, $SD = 0.57$) than on target-absent line-up judgements ($M = 2.42$, $SD = 0.60$), leading to a main effect of line-up type, $F(1, 20) = 43.64$, $p < .001$, $\eta_p^2 = .69$. The interaction was not significant, $F(1, 20) < 1$.

A 2 (recognition accuracy: incorrect vs correct) $\times$ 2 (awareness: change blindness vs change detection) repeated-measures ANOVA was conducted with confidence as the dependent variable. This test revealed main effects of accuracy, $F(1, 20) = 96.73$, $p < .001$, $\eta_p^2 = .83$, and awareness, $F(1, 20) = 11.62$, $p = .003$, $\eta_p^2 = .37$, with no interaction, $F(1, 20) < 1$. Paired samples *t*-tests showed confidence was higher for correct recognitions than for incorrect recognitions after change blindness, $t(20) = -4.71$, $p < .001$, $d = 1.11$, as well as after change detection, $t(20) = -8.60$, $p < .001$, $d = 1.94$ (see Table 1).

Discussion

The results of Experiment 3 largely replicated those reported in Experiments 1 and 2. Again, recognition confidence was higher after change detection than after change blindness. Because faces were not repeated and no feedback about the change detection task was given, the possibility that these factors were confounding the results from the first two experiments can be discounted. In addition to change detection performance, confidence was strongly affected by recognition accuracy. Confidence was substantially higher for correct recognition choices than for incorrect recognition choices. Although it is clear that accuracy accounted for some of the variance in confidence for the target-present trials, the same cannot be said for the target-absent trials. For the trials on which it was impossible to make an accurate recognition choice, confidence was higher when the change was detected than when it was not. Similar to Experiment 1, confidence was higher for correct recognitions than for incorrect recognitions on both change detection and change blindness trials. Consistent with the optimality hypothesis, the effect size was larger for change detection trials than for change blindness trials, although both effect sizes were large according to Cohen's (1988) classifications.

EXPERIMENT 4

In the first three experiments participants were forced to choose one of the faces from the line-up. This would not be the case in an actual line-up judgement task, in which eyewitnesses have the option of rejecting the line-up if they believe the perpetrator to be absent. Without the option to reject the line-up, participants in our experiments may be expressing their confidence *given the difficulty of the choice*, rather than their confidence in their memory for the displays per se. Thus we conducted an additional experiment in which participants were given the opportunity to reject the line-up altogether. One advantage of this design is that it provides the opportunity to examine the influence of choosing on confidence and accuracy. This variable is of particular interest because, according to Sporer et al.'s (1995) meta-analysis, the confidence–accuracy relation is stronger when a line-up member is chosen than when the line-up is rejected.

In addition to analysing the effect of choosing on the confidence–accuracy relation, we also explored how detection of change influences choosing behaviour. Previous research has shown choosing is influenced by a host of factors, including the line-up procedure that is used (Gronlund, Carlson, Dailey, & Goodsell, 2009; Memon & Gabbert, 2003), whether or not line-up instructions are biased (Steblay, 1997), and the physical characteristics of the line-up members (Wilcock, Bull, & Vrij, 2007). In a previous investigation that included a manipulation of encoding conditions, the distance between the witness and the target had no influence on choosing behaviour (Lindsay, Semmler, Weber, Brewer, & Lindsay, 2008). However, in the present study poor encoding likely results from inattention to the location of the change, which might propagate onto the recognition memory judgement. After detecting the change, participants might be more willing to choose someone from the line-up based on a feeling of overconfidence stemming from the perception of good encoding conditions. Alternatively, the perception of good encoding conditions after change detection might lead participants to require a stronger signal of memory strength and lead to a more conservative response bias.

Method

Participants. A total of 20 university undergraduates participated in exchange for either partial course credit or payment of $10.

Materials. The stimuli set consisted of 342 faces. All faces from Experiment 3 were used as well as nine additional faces that were taken from the Center for Vital Longevity Database. As in Experiment 3, faces were not repeated across trials.

Procedure. The procedures were the same as in Experiment 3, with the exception that participants were given the option to reject line-ups and the proportion of target-absent line-ups was increased from 25% to 50%.

Results

The small number of trials precluded carrying out analyses involving all three factors (line-up type, awareness, and line-up decision), so the data were analysed with three separate two-factor ANOVAs instead. A 2 (line-up type: target-present vs target-absent) $\times$ 2 (awareness: change blindness vs change detection) repeated-measures ANOVA revealed higher confidence for change detection trials than for change blindness trials (see Table 1), yielding a significant main effect of awareness, $F(1, 18) = 16.33$, $p = .001$, $\eta_p^2 = .48$. Paired samples *t*-tests confirmed confidence was higher on change detection trials than on change blindness trials for both target-present trials, $t(18) = -3.47$, $p = .003$, $d = .79$, and for target-absent trials, $t(19) = -3.72$, $p = .001$, $d = .84$. In addition, confidence was higher on target-present line-ups ($M = 3.40$, $SD = 0.66$) than on target-absent line-ups ($M = 3.02$, $SD = 0.56$), yielding a main effect of line-up type, $F(1, 18) = 8.51$, $p = .009$, $\eta_p^2 = .32$. The interaction was not significant, $F(1, 18) = 1.60$, $p = .22$, $\eta_p^2 = .08$.

A 2 (recognition accuracy: incorrect vs correct) $\times$ 2 (awareness: change blindness vs change detection) repeated-measures ANOVA on confidence revealed a significant interaction, $F(1, 18) = 4.56$, $p = .047$, $\eta_p^2 = .20$. The interaction was indicative of a greater effect of awareness on confidence for correct recognitions than for incorrect recognitions. Paired samples *t*-tests showed higher confidence for correct recognitions than for incorrect recognitions on change blindness trials, $t(19) = -3.13$, $p = .006$, $d = 0.69$, and change detection trials, $t(18) = -4.91$, $p < .001$, $d = 1.12$ (see Table 1).

A 2 (recognition accuracy: incorrect vs correct) $\times$ 2 (line-up decision: choose vs reject) repeated-measures ANOVA on confidence revealed main effects of accuracy, $F(1, 16) = 31.65$, $p < .001$, $\eta_p^2 = .66$, and line-up decision, $F(1, 16) = 22.95$, $p < .001$, $\eta_p^2 = .59$, with no interaction, $F(1, 16) < 1$. Paired samples *t*-tests showed correct recognitions were made more confidently than incorrect recognitions for choosing trials, $t(19) = -7.24$, $p < .001$, $d = 1.65$, as well as for rejection trials, $t(16) = -2.85$, $p = .01$, $d = 0.76$.

A paired samples *t*-test was conducted to examine whether or not awareness of change influenced the probability of choosing someone from the line-up. This test showed no reliable difference in choosing between change detection ($M = 0.76$, $SD = 0.16$) and change blindness ($M = 0.70$, $SD = 0.18$) trials, $t(19) = -1.32$, $p = .20$, $d = 0.30$. Therefore we found no indication that encoding conditions influenced choosing behaviour.

Discussion

Once again, confidence was higher after change detection than after change blindness. This effect was consistent regardless of line-up type and recognition accuracy. The effect of awareness on the confidence–accuracy relation was also replicated: Consistent with the optimality hypothesis, the strength of association was larger for change detection trials than for change blindness trials.

The effect of choosing on the confidence–accuracy relation was consistent with the results of Sporer et al.'s (1995) meta-analysis. Although Sporer et al. reported higher confidence for correct identifications than for incorrect identifications for both choosers and non-choosers, the mean difference was substantially larger for choosers. Similarly, correct recognitions in the present research were made more confidently than incorrect recognitions for trials on which faces were chosen as well as for trials on which line-ups were rejected. Furthermore, the effect size for choosing trials was more than double the effect size for rejection trials. Thus the effect of choosing corresponds well with previous research.

We also examined how awareness of change influenced the likelihood that participants would choose someone from the line-up. This analysis showed the probability of choosing was similar for change detection and change blindness trials, suggesting awareness of change did not influence choosing behaviour. This is consistent with the findings of Lindsay et al. (2008), who examined the role of a different form of encoding conditions (i.e., witness-perpetrator distance) on choosing tendencies.

META-COMPARISON OF EXPERIMENTS 1, 3, AND 4

In the foregoing analyses the confidence–accuracy relation consisted of the mean difference in confidence ratings between correct and incorrect recognitions. One advantage of this approach is that it provides a powerful statistical test of the confidence–accuracy relation (Robinson & Johnson, 1998). For a more precise indication of the confidence–accuracy relation, however, we calculated two measures of resolution (i.e., how effectively confidence ratings correspond to correct and incorrect recognitions). One measure, the Goodman-Kruskal gamma coefficient (G), was originally recommended by Nelson (1984, 1986) and is commonly used in metamemory research. However, gamma can be influenced by response bias (Masson & Rotello, 2009), and simulation analyses have shown that it does not perform as well as measures that are based on signal detection theory (Rotello, Masson, & Verde, 2008). Thus we also calculated the area under the receiver operating characteristic curve (A_z), an alternative to gamma suggested by Masson and Rotello (2009).

Results and discussion

Average scores for the gamma coefficient and for the area under the receiver operating characteristic curve were computed for each participant from Experiments 1, 3, and 4. One sample t-tests revealed positive confidence–accuracy relations. The average of all participants' average gamma coefficient ($M = .41$, $SD = .26$) was significantly higher than 0, $t(65) = 12.96$, $p < .001$. Similarly, the average of all participants' average area under the receiver operating characteristic curve ($M = .68$, $SD = .11$) was significantly higher than 0.5, $t(65) = 13.26$, $p < .001$.

Both measures indicated awareness of change had an influence on the confidence–accuracy relation. For gamma coefficients, a paired samples t-test revealed significantly better resolution on change detection trials ($M = .45$, $SD = .29$) than on change blindness trials ($M = .30$, $SD = .39$), $t(65) = -2.58$, $p = .01$, $d = .31$. Similarly, for areas under the receiver operating characteristic curve, the difference between change detection trials ($M = .69$, $SD = .15$) and change blindness trials ($M = .63$, $SD = .20$) was significant, $t(65) = -2.63$, $p = .01$, $d = 0.30$. Thus these metamemory measures converged with the comparisons of confidence on correct and incorrect trials reported within each experiment separately in providing support for the optimality hypothesis.

GENERAL DISCUSSION

In four experiments participants sequentially viewed two arrays of faces. Across the two arrays, five of the faces were consistent and one of the faces changed. Their task was to locate the face that changed, recognise one of the previously viewed faces from a line-up, and then rate their recognition confidence. According to Deffenbacher's (1980) optimality hypothesis, good encoding conditions should facilitate a strong confidence–accuracy relation. Consistent with the optimality hypothesis we found stronger confidence–accuracy relations when the change was detected compared to when it was not detected (i.e., change blindness). However, not all of our findings were consistent with the optimality hypothesis. Specifically, in Experiment 2 and the target-absent trials of Experiment 3, confidence ratings were higher on change detection trials than on change blindness trials, even though none of the recognitions was accurate. In contrast to the optimality hypothesis, which predicts confidence and accuracy will be affected by encoding conditions uniformly, we observed an inflated sense of confidence on change detection trials when accurate recognitions were impossible.

Participants in our study were not directly asked whether their confidence judgements were related to their perception of how well the displays were encoded. Nevertheless, our findings are consistent with previous claims that participants rely on a meta-memorial judgement of encoding conditions to evaluate their confidence in their choice on a recognition memory task (Lindsay et al., 1998). Participants might make an immediate appraisal of the encoding conditions, and use this to determine both the criterion they will establish for judging a line-up member as ''old'' and for evaluating the accuracy of that judgement. Bruno, Higham, and Perfect (2009) predicted that participants would apply a similar strategy on a word-learning task. They postulated a global subjective memorability hypothesis wherein participants would adjust their response criterion according to how they perceived the learning conditions. This hypothesis was supported when, in a series of experiments, participants who judged learning conditions as poor relied on extra-memorial cues to guide their recognition decisions. In contrast, when learning conditions were good, participants based their criterion on memory strength. In the present experiments participants took the learning conditions into account both when deciding if a signal was present as well as when indicating their level of certainty in that decision.

Why does change detection inflate confidence on a subsequent task? The answer to this question might be that succeeding in detecting a change functions as feedback about the quality of the encoding conditions. In Experiments 1 and 2, participants were explicitly told whether or not they correctly localised a change they knew with certainty had occurred. In Experiments 3 and 4, although we did not provide any external feedback, the act of detecting the change might have acted as a source of internal feedback.

Previous researchers have found that pre-identification feedback about encoding conditions can influence identification confidence. In one study (Gabbert, Memon, & Wright, 2007) participants who received feedback suggesting that the conditions in which encoding occurred were favourable were more confident in their memory than were participants who received feedback suggesting that the same encoding conditions were unfavourable. Similar results have been found by Leippe and colleagues (Leippe, Eisenstadt, & Rauch, 2009; Leippe, Eisenstadt, Rauch, & Stambush, 2006), who manipulated feedback about memory reports and found perceptions of memory accuracy influenced line-up identification confidence. Although the reliability of their results differed depending on how feedback was manipulated, positive feedback generally led to higher confidence than negative or no feedback. Similarly, in the present research participants who received positive feedback about the change detection task were more confident than participants who received negative feedback.

For some time now researchers have warned of the dangers of line-up administrators providing witnesses with feedback about the veracity of their identification (Bradfield, Wells, & Olson, 2002; Wells & Bradfield, 1998). In contrast to the well-documented problems associated with post-identification feedback, our research casts light on the less well-understood problem of pre-identification feedback. In a real-world case such feedback could consist of information provided to a witness during the initial interview with a police investigator. For example, an eyewitness could report observing the perpetrator and two other witnesses, and the officer could inadvertently confirm that claim by mentioning that the other two witnesses have spoken with police. Our findings suggest positive feedback of this nature could falsely increase confidence in a subsequent line-up identification.

The relation between confidence and accuracy is not well understood, yet jurors use confidence as a means of assessing the credibility of eyewitness identifications (Brewer & Burke, 2002). Therefore it is critical to study the factors that influence the relation between these two variables. Our findings imply that recognition confidence is not always based on the strength of the memory representation upon which the recognition judgement relies. To the contrary, we found evidence to suggest that participants integrated information about how well they encoded a scene with their actual memory for the scene when they determined their recognition confidence. The results underscore the importance of assessing the source of confidence in tests of recognition memory, suggesting high confidence may reflect the perceived quality of the conditions in which a scene was encoded, rather than the strength of the memory for the scene itself.

Manuscript received 31 October 2010
Manuscript accepted 20 July 2011
First published online 21 October 2011

REFERENCES

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*, 691–695.

Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology*, *87*, 112–120. doi: 10.1037//0021-9010.87.1.112

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior*, *26*, 353–364.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, *7*, 207–218. doi: 10.1037//1076-898X.7.3.207

Bruno, D., Higham, P. A., & Perfect, T. J. (2009). Global subjective memorability and the strength-based mirror effect in recognition memory. *Memory & Cognition*, *37*, 807–818. doi: 10.3758/MC.37.6.807

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davies, G., & Hine, S. (2007). Change blindness and eyewitness testimony. *Journal of Psychology*, *141*, 423–434.

Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law and Human Behavior*, *4*, 243–260.

Gabbert, F., Memon, A., & Wright, D. B. (2007). I saw it for longer than you: The relationship between perceived encoding duration and memory conformity. *Acta Psychologica*, *124*, 319–331. doi: 10.1016/j.actpsy.2006.03.009

Gronlund, S. D., Carlson, C. A., Dailey, S.B., & Goodsell, C. A. (2009). Robustness of the sequential line-up advantage. *Journal of Experimental Psychology: Applied*, *15*, 140–152. doi: 10.1037/a0015082

Leippe, M. R., Eisenstadt, D., & Rauch, S. M. (2009). Cueing confidence in eyewitness identifications: Influence of biased line-up instructions and pre-identification memory feedback under varying line-up conditions. *Law and Human Behavior*, *33*, 194–212. doi: 10.1007/s10979-008-9135-y

Leippe, M. R., Eisenstadt, D., Rauch, S. M., & Stambush, M. A. (2006). Effects of social-comparative memory feedback on eyewitnesses' identification confidence, suggestibility, and retrospective memory reports. *Basic and Applied Social Psychology*, *28*, 201–220. doi: 10.1207/s15324834basp2803_1

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*, 501–506.

Levin, D. T., Simons, D. J., Angelone, B. L., & Chabris, C. F. (2002). Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*, *93*, 289–302.

Lindsay, D. S., Nilsen, E., & Read, J. D. (2000). Witnessing-condition heterogeneity and witnesses' versus investigators' confidence in the accuracy of witnesses' identification decisions. *Law and Human Behavior*, *24*, 685–687. doi: 0147-7307/00/1200-0685

Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*, 215–218.

Lindsay, R. C. L., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. R. (2008). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behaviour*, *32*, 526–535. doi: 10.1007/s10979-008-9128-x

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 509–527. doi: 10.1037/a0014876

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*, 865–876.

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, *69*, 1175–1184.

Memon, A., & Gabbert, F. (2003). Improving the identification accuracy of senior witnesses: Do pre-line-up questions and sequential testing help? *Journal of Applied Psychology*, *88*, 341–347. doi: 10.1037/0021-9010.88.2.341

Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, *94*, 339–354.

Minear, M., & Park, D.C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments & Computers*, *36*, 630–633.

Nelson, K. J., Laney, C., Bowman Fowler, K., Knowles E. D., Davis, D., & Loftus, E. F. (2011). Change blindness can cause mistaken eyewitness identification. *Legal and Criminological Psychology*, *16*, 62–74. doi: 10.1348/135532509X482625

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133.

Nelson, T. O. (1986). BASIC programs for computation of the Goodman-Kruskal gamma coefficient. *Bulletin of the Psychonomic Society*, *24*, 281–283.

Read, J. D. (1995). The availability heuristic in person identification: The sometimes misleading consequences of enhanced contextual information. *Applied Cognitive Psychology*, *9*, 91–121.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across brief interruptions. *Visual Cognition*, *7*, 127–145.

Robinson, M. D., & Johnson, J. T. (1998). How not to enhance the confidence–accuracy relation: The detrimental effects of attention to the identification process. *Law and Human Behavior*, *22*, 409–428.

Rotello, C. M., Masson, M. E. J., & Verde, M. F. (2008). Type I error rates and power analyses for single-point sensitivity measures. *Perception & Psychophysics*, *70*, 389–401. doi: 10.3758/PP.70.2.389

Simons, D. J., & Levin, D. T (1997). Change blindness. *Trends in Cognitive Sciences*, *1*, 261–267. doi: 10.1016/S1364-6613(97)01080-2

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence accuracy relation in eyewitness studies. *Psychological Bulletin*, *118*, 315–327.

Steblay, N. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, *21*, 283–297.

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, *83*, 360–376.

Wilcock, R. A., Bull, R., & Vrij, A. (2007). Are old witnesses always poorer witnesses? Identification accuracy, context reinstatement, own-age bias. *Psychology, Crime & Law*, *13*, 305–316. doi: 10.1080/10683160600822212