G Model
JARMAC-171; No. of Pages 8
ARTICLE IN PRESS

Journal of Applied Research in Memory and Cognition xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Original Article

Effects of eye-closure on confidence-accuracy relations in eyewitness testimony

Annelies Vredeveldt[a,*], James D. Sauer[b]

[a] Department of Criminal Law and Criminology, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
[b] Department of Psychology, University of Tasmania, Australia

ARTICLE INFO

Article history:
Received 29 January 2014
Received in revised form
13 December 2014
Accepted 24 December 2014
Available online xxx

Keywords:
Eyewitness memory
Metacognitive monitoring
Eye-closure
Confidence-accuracy calibration
Overconfidence
Precision

ABSTRACT

Confidence judgments about the quality of memory can have serious implications in eyewitness settings. Three experiments investigated the effect of eye-closure during eyewitness interviews on confidence-accuracy relations in event recall. In all experiments, participants viewed video-taped events and were subsequently questioned about the event, while they had their eyes open or closed. Participants provided confidence ratings for each response. We found that participants were generally able to monitor the accuracy of their responses, although they displayed underconfidence for imprecise responses. Importantly, across all experiments, eye-closure increased accuracy without significantly inflating confidence or impairing confidence-accuracy relations. Moreover, in Experiment 3, reducing distraction (e.g., through eye-closure) significantly reduced overconfidence. Thus, unlike most other investigative interview protocols that facilitate recall, eye-closure improves recall accuracy with no apparent cost, and some evidence of benefit, to metamemory. Practical implications of these findings are discussed, and hypotheses regarding potential theoretical mechanisms are proposed.

Confidence is fundamental to the regulation of memory reporting. It determines whether recalled information is volunteered or withheld, and the level of detail reported (Goldsmith, Koriat, & Weinberg-Eliezer, 2002; Koriat & Goldsmith, 1996; Weber & Brewer, 2008). Furthermore, confidence expressed by eyewitnesses (e.g., "I am 100% certain he had a gun") influences investigators', jurors', and judges' assessments of witnesses' reliability. Thus, it is important to consider not only how interviewing methods affect recall, but also how they affect witness confidence, and witnesses' ability to discriminate between information that is more or less likely to be reliable. We examined whether eye-closure, a method that facilitates event recall (e.g., Perfect et al., 2008; Vredeveldt & Penrod, 2013), affects the confidence-accuracy (CA) relationship in eyewitness memory.

Witness confidence can affect criminal investigations. Police are likely to place greater weight on, and devote greater investigative resources to pursuing, details about which eyewitnesses are certain. Witness confidence is also influential in court. Expressions of confidence may help judges and jurors decide whether a particular detail is accurate, and discriminate between witnesses (or details) that are more or less likely to be accurate. When a witness appears confident, jury-eligible samples and legal professionals are more likely to believe that the witness is accurate, and the defendant guilty (Brewer & Burke, 2002; Brigham & Wolfskeil, 1983; Cutler, Penrod, & Dexter, 1990; Noon & Hollin, 1987).

However, the diagnostic value of confidence depends on individuals' ability to accurately evaluate their own memory. In their seminal article, Nisbett and Wilson (1977) observed that individuals often lack introspective access into higher-order cognitive processes. Metacognitive judgments tend to rely on inferential processes (Koriat, 1993, 2012), and can be (a) distorted by various non-memorial influences and (b) insensitive to variations in memory quality, impairing individuals' ability to discriminate correctly from incorrectly recalled details. Further, people often overestimate the reliability of recalled information (i.e., display overconfidence; Fischhoff, Slovic, & Lichtenstein, 1977; Gigerenzer, Hoffrage, & Kleinbölting, 1991; Koriat, Lichtenstein, & Fischhoff, 1980). When witnesses assess confidence in their testimony as a whole, their global confidence rating typically does not correlate significantly with recall accuracy (Granhag, 1997; Granhag, Jonsson, & Allwood, 2004; Gwyer & Clifford, 1997; Mello & Fisher, 1996; Wagstaff et al., 2004). In contrast, when witnesses provide separate confidence ratings for each response, confidence and accuracy tend to be positively correlated (Allwood, Ask, & Granhag, 2005; Roberts & Higham, 2002; Wagstaff et al., 2004).

* Corresponding author. Tel.: +31 205983153.
E-mail address: anneliesvredeveldt@gmail.com (A. Vredeveldt).

http://dx.doi.org/10.1016/j.jarmac.2014.12.006
2211-3681/

Please cite this article in press as: Vredeveldt, A., & Sauer, J.D. Effects of eye-closure on confidence-accuracy relations in eyewitness testimony. *Journal of Applied Research in Memory and Cognition* (2015), http://dx.doi.org/10.1016/j.jarmac.2014.12.006

Several methods that facilitate eyewitness memory have unwanted effects on confidence. Hypnosis typically increases the amount of information reported, but also consistently inflates confidence in false memories (Dywan & Bowers, 1983; Kebbell & Wagstaff, 1998). Similarly, mental context reinstatement increases remembering but can also inflate confidence (Hammond, Wagstaff, & Cole, 2006). More complicated findings have been reported for the Cognitive Interview, an interviewing protocol that incorporates various rapport-building and mnemonic techniques to enhance recall (see Fisher & Geiselman, 1992). Some research suggests that the Cognitive Interview does not significantly affect certain indices of confidence, such as overall confidence (e.g., McCauley & Fisher, 1995; McMahon, 2000), confidence in erroneous recall (e.g., Granhag et al., 2004), and CA correlations (e.g., Gwyer & Clifford, 1997). However, Allwood et al. (2005) found that the Cognitive Interview decreased discrimination between accurate and inaccurate responses, and Granhag et al. (2004) found that it increased overconfidence.

Recently, researchers have proposed a simple method to facilitate remembering: closing the eyes during recall. This method increases both the amount and the accuracy of event recall (e.g., Perfect et al., 2008; Vredeveldt, Hitch, & Baddeley, 2011; Wagstaff et al., 2004). Most previous investigations of the eye-closure effect have not reported measures of confidence, with the exception of Wagstaff and colleagues. Across three studies (Wagstaff et al., 2004; Wagstaff, Wheatcroft, Burt, et al., 2011; Wagstaff, Wheatcroft, Caddick, Kirby, & Lamont, 2011), they found no significant effect of eye-closure on mean confidence in correct and incorrect responses about witnessed events. Additionally, Wagstaff et al. (2004) found a non-significantly higher correlation between accuracy and confidence for participants who closed their eyes ($r = .74$), compared to participants who kept their eyes open ($r = .57$). In sum, limited findings to date suggest that eye-closure has no or minimal effects on the CA relation.

We extend previous work on the effect of eye-closure on the CA relation, using more comprehensive and sensitive analyses. First, researchers often fail to take into account that individuals regulate the precision of their answers to compensate for reductions in memory quality (Goldsmith et al., 2002; Goldsmith, Koriat, & Pansky, 2005), which may conceal important effects on memory output and the CA relation (see also Fisher, 1996). Therefore, we conducted separate analyses for precise and imprecise responses. Second, because the informative value of point-biserial correlations is limited (i.e., robust CA relationships are compatible with correlations ranging from near-zero to 1; Juslin, Olsson, & Winman, 1996), we inspected a range of other confidence measures. Across three experiments, we examined the effect of eye-closure on recall accuracy, mean confidence, and discrimination between accurate and inaccurate responses (measured by adjusted normalized discrimination index [ANDI]; e.g., Yaniv, Yates, & Smith, 1991). Further, our 0–100% confidence scale in Experiment 3 permitted calculation of calibration and over/underconfidence statistics (which will be explained in more detail under Experiment 3).

1. Experiment 1

1.1. Method

1.1.1. Participants

Fifty-six students participated for course credit or a small monetary reward (11 male and 45 female; mean age = 19.91, SD = 2.47).

1.1.2. Materials

One violent and one non-violent version were created for two episodes of different TV shows, resulting in four 8-min video clips. The first episode was about survivors of a plane crash on an apparently deserted island, who discover a house. The second was about a woman looking for her missing son in a forest. The violent versions for each episode included a gun or arrow shot, stitching up of a wound, and a physical fight, whereas the non-violent versions showed explorations of the house and peaceful interactions. For each version of the video, a set of twenty questions was constructed, addressing visual (e.g., "Where on his body does the man get shot?") and auditory (e.g., "Where does the man say that the medical kit is?") aspects of the events. Questions were asked in chronological order.

1.1.3. Design

Interview condition (eyes open or closed) and type of event (violent or non-violent) were manipulated between- and within-subjects, respectively. Participants were randomly assigned to condition. Participants watched two videos: the violent version of one TV show and the non-violent version of the other show, with the order of videos counterbalanced.[1]

1.1.4. Procedure

Participants provided informed consent, watched the first video, completed a two-minute filler task (a word finder), and responded orally to questions about the first video. Depending on condition, participants were either instructed to keep their eyes closed throughout the interview (and reminded appropriately), or received no instruction. Participants were instructed to answer questions in as much detail as possible, but not to guess: A "don't know" response was allowed. After each response, participants indicated their confidence on a scale of 1 ("not confident at all") to 5 ("extremely confident"). This procedure was repeated for the second video. Interviews were audio-taped.

1.1.5. Data coding

Interviews were coded blind to condition. Responses were coded as correct, incorrect, or omitted ("don't know"). We employed a relatively strict scoring procedure, in which a response was scored as incorrect if it contained any incorrect elements, even if part of the answer was accurate. Responses were also coded for precision, or the level of specificity provided.[2] For example, in response to the question "Where on his body does the man get shot?", possible answers could be "on his left upper arm" (correct, precise), "on his arm" (correct, imprecise), "on his right upper arm" (incorrect, precise), "on his leg" (incorrect, imprecise), or "don't know" (omitted). For each of the four video clips, the responses of five randomly selected participants were double-coded by an independent coder (i.e., 100 responses per video; 400 responses in total; 18% of the total sample). Interrater reliability was high, $\kappa = .93$, $p < .001$. The codes of the first coder were retained for analysis.

1.2. Results

1.2.1. Data transformations

Prior to all analyses reported in this article, relevant assumptions were checked. Where appropriate, skewness was countered through square-root transformations. Descriptive statistics are based on the untransformed variables.

[1] For the present purposes, we focus only on the effect of eye-closure on the confidence-accuracy relation, which was not affected by type of video, presentation order, or question modality.

[2] This concept is akin to the concept of "grain size" proposed by Goldsmith et al. (2002), except that grain size generally refers to the specificity of a single descriptive element (e.g., "brown" versus "mahogany"), whereas our definition of precision refers to the specificity of the answer as a whole (e.g., "brown" versus "brown and curly" hair).

1.2.2. Recall accuracy

Proportion correct was calculated by dividing the number of correct responses by the number of correct plus incorrect responses. A 2 (Interview Condition: eyes open, eyes closed) by 2 (Precision: precise, imprecise) mixed ANOVA[3] on proportion correct revealed that eye-closure significantly increased accuracy, F (1, 54) = 5.98, p = .018, f = .33, and that imprecise answers were significantly more accurate than precise answers, F (1, 54) = 8.70, p = .005, f = .27. There was also a significant interaction between condition and precision, F (1, 54) = 5.38, p = .024, f = .22. Eye-closure significantly increased accuracy for precise details (eyes-open: M = .79, SD = .11; eyes-closed: M = .90, SD = .08), F (1, 54) = 19.71, p < .001, f = .60, but did not significantly affect accuracy for imprecise details (eyes-open: M = .88, SD = .10; eyes-closed: M = .89, SD = .09), F (1, 54) = .03, p = .871, f = .02.

To examine whether eye-closure affected participants' response criterion, we conducted an independent t-test on the mean number of "don't know" responses in each condition, which revealed no significant difference between the eyes-open (M = 2.14, SD = 1.73) and eyes-closed (M = 1.98, SD = .91) conditions, t (40.73) = .33, p = .747, f = .06.

1.2.3. Mean confidence

Confidence ratings were nested within participants (i.e., each participant provided a different number of correct and incorrect, and precise and imprecise, responses). To account for the effects of accuracy and precision, we conducted multilevel modeling on confidence ratings (cf. Peugh & Enders, 2005). The unconditional means model with participant as a Level 2 variable (Model A) indicated significant variability in confidence ratings, both within-participants (τ^{00} = .41) and between-participants (σ^2 = .06). In Model B, we added our main predictor (interview condition) as a between-participants fixed effect, which did not significantly improve the fit of the model, $\chi^2(1)$ = 0.40, p = .525, PRV = .01. In Model C, we added two within-participants covariates (accuracy and precision), which significantly improved model fit, $\chi^2(2)$ = 456.19, p < .001, PRV = .21. In Model D, we added all potential two- and three-way interactions, but this did not significantly improve model fit, $\chi^2(4)$ = 8.02, p = .091, PRV = .00, and none of the interactions were significant. Therefore, we selected Model C as our final model, which reduced between-subjects variability by 7.7% and within-subjects variability by 20.3% compared to the unconditional means model.

Model C revealed a small, non-significant difference in confidence ratings between the eyes-open condition (M = 4.03, SD = 1.12) and the eyes-closed condition (M = 3.97, SD = 1.12), F (1, 56.69) = 2.84, p = .097, f = .22. Participants were significantly more confident in accurate responses (M = 4.19, SD = .97) than inaccurate responses (M = 2.83, SD = 1.21), F (1, 1966.37) = 416.49, p < .001, f = .46, and significantly more confident in precise (M = 4.14, SD = 1.04) than imprecise responses (M = 3.67, SD = 1.22), F (1, 1970.71) = 95.29, p < .001, f = .22.

1.2.4. Discrimination between accurate and inaccurate responses

The adjusted normalized discrimination index (ANDI) indicates the extent to which a participant's confidence ratings discriminate correct from incorrect responses (for a detailed account of the formulae for calculating this index, see Yaniv et al., 1991). ANDI ranges from 0 (no discrimination) to 1 (perfect discrimination), and reflects the amount of variance in accuracy accounted for by participants' confidence ratings. Yaniv and colleagues note that ANDI has two benefits over previous measures of discrimination skill: (1) interpretation of ANDI is not conditional on the objective uncertainty of the outcome being predicted, thus removing a "statistical ceiling" from the achievable level of discrimination, and (2) ANDI is unaffected by the number of judgment categories (e.g., the use of 1–5 scale vs. the use of a 0–100 scale), or the number of judgment cases.[4]

A two-way ANOVA on ANDI revealed no significant effects of condition, F (1, 34) = 0.83, p = .370, f = .16, or precision, F (1, 34) = 2.09, p = .157, f = .16, and no significant interaction, F (1, 34) = 1.28, p = .266, f = .16. Mean ANDI for precise responses was .41 (SD = .29) in the eyes-open condition and .37 (SD = .31) in the eyes-closed condition. Mean ANDI for imprecise responses was .27 (SD = .33) in the eyes-open condition and .41 (SD = .39) in the eyes-closed condition.

1.3. Discussion

Experiment 1 confirmed previous findings that eye-closure improves recall accuracy (e.g., Perfect et al., 2008). Eye-closure improved accuracy for precise, but not imprecise responses. Eye-closure did not significantly affect mean confidence in responses or participants' ability to discriminate between correct and incorrect responses. To test the generalizability of this pattern, the next two studies examined confidence measures collected during previous work on the eye-closure effect (Experiment 2: Vredeveldt, Baddeley, & Hitch, 2014; Experiment 3: Vredeveldt et al., 2011).[5]

2. Experiment 2

Vredeveldt and Baddeley (2014a) examined the effects of eye-closure in a repeated-recall paradigm. Here, we analyze the confidence ratings provided during cued recall in that study to investigate whether the pattern reported in Experiment 1 holds when the cued-recall interview is preceded by a one-week delay and two previous free-recall attempts. Two lines of research highlight this as an important area for investigation. First, answering the same question repeatedly can inflate confidence (e.g., Odinot, Wolters, & Lavender, 2009; Shaw, 1996; Shaw & McClure, 1996; but see Odinot, Wolters, & van Giezen, 2013). Second, confidence can be relatively insensitive to factors that reduce memory quality (e.g., increased retention interval), again leading to increased overconfidence (e.g., Odinot & Wolters, 2006; Sauer, Brewer, Zweck, & Weber, 2010). Thus, we explored how eye-closure affected CA relations under conditions that promote overconfidence.

The introduction of a one-week delay and multiple recall attempts also enhanced the ecological validity of the research. In applied settings, the delay between experiencing an event and recalling the event will typically be longer than a few minutes, and individuals may attempt to retrieve a particular event on multiple occasions (e.g., during repeated witness interviews). Thus, an additional aim of Experiment 2 was to investigate the impact of eye-closure on CA relations under more realistic conditions.

[3] Although modality of recalled details was not our main focus, we report the findings here for interested readers. A 2 (Interview Condition: eyes open, eyes closed) by 2 (Modality: visual, auditory) mixed ANOVA on proportion correct revealed no significant interaction between eye-closure and modality, F (1, 54) = 0.55, p = .461, f = .06. Pairwise comparisons showed that eye-closure significantly increased accuracy for visual information (eyes-open: M = .84, SD = .09; eyes-closed: M = .93, SD = .06), F (1, 54) = 15.01, p < .001, f = .53, as well as auditory information (eyes-open: M = .80, SD = .09; eyes-closed: M = .87, SD = .09), F (1, 54) = 9.90, p = .003, f = .43.

[4] Although the Goodman-Kruskal Gamma coefficient (Goodman & Kruskal, 1954) is a relatively common index of metacognitive (discrimination) skill, it is systematically distorted by variations in response bias (that are independent of discrimination skill; see Masson & Rotello, 2009).

[5] Confidence measures were not analyzed in previous publications.

2.1. Method

A detailed description of the methodology is provided by Vredeveldt, Baddeley, and Hitch (2014). In brief, 48 participants were randomly assigned to one of four combinations of instructed eye-closure during interviews in sessions one and two (open-open, closed-open, open-closed, or closed-closed). In the first session, participants watched a video depicting a violent encounter between a man and a woman, completed a two-minute distracter task, and provided a free recall of the event. In the second session one week later, participants first provided another free recall, and then participated in a cued-recall interview (with 16 questions about the event). After each cued-recall response, participants indicated their confidence on a scale from 1 ("not confident at all") to 5 ("extremely confident"). Because confidence ratings were obtained only for cued-recall responses, the free-recall data will not be discussed further (but see Vredeveldt, Baddeley, & Hitch, 2014). For the present research, audio-taped interviews were re-coded using the coding procedure described for Experiment 1. Ten interviews (160 responses; 21% of the total sample) were randomly selected and scored independently by a second blind coder. Interrater reliability was high, $\kappa = .83$, $p < .001$. The codes of the first coder were retained for the main analysis.

2.2. Results

2.2.1. Recall accuracy

Data transformations on proportion correct did not attenuate problems with negative skewness and heterogeneity of variance, hence we conducted non-parametric tests. To enable comparisons with Experiments 1 and 3, we also provide means, standard deviations, and Cohen's f effect sizes. A Mann–Whitney test revealed that participants who closed their eyes were again significantly more accurate ($M = .76$, SD = .11) than participants who kept their eyes open ($M = .67$, SD = .13), $U = 180.50$, $p = .026$, $f = .38$, and a Wilcoxon signed-rank test showed that imprecise answers ($M = .78$, SD = .21) were significantly more accurate than precise answers ($M = .68$, SD = .18), $T = 311.50$, $p = .031$, $f = .27$. Interview condition again had no significant effect on accuracy for imprecise details, $U = 572.50$, $p = .749$, $f = .04$, but unlike in Experiment 1, the effect of condition on accuracy for precise details was also non-significant, $U = 499.00$, $p = .066$, $f = .30$ (although the effect was moderate in size). Finally, there was no significant difference in the number of "don't know" responses between the eyes-open ($M = 4.71$, SD = 1.94) and eyes-closed ($M = 3.96$, SD = 1.81) conditions, $t(46) = 1.38$, $p = .173$, $f = .20$.

2.2.2. Mean confidence

As in Experiment 1, we conducted multilevel modeling on confidence ratings, and Model C best fit the data (reducing between-subjects variability by 26.0% and within-subjects variability by 13.0%). Again, there was no significant difference between the eyes-open ($M = 3.11$, SD = 1.27) and eyes-closed conditions ($M = 3.22$, SD = 1.26), $F(1, 48.28) = .07$, $p = .799$, $f = .04$. Participants were significantly more confident in accurate ($M = 3.42$, SD = 1.21) than inaccurate responses ($M = 2.52$, SD = 1.17), $F(1, 541.32) = 71.21$, $p < .001$, $f = .36$, and significantly more confident in precise ($M = 3.34$, SD = 1.27) than imprecise responses ($M = 2.90$, SD = 1.22), $F(1, 544.09) = 23.90$, $p < .001$, $f = .21$.

2.2.3. Discrimination between accurate and inaccurate responses

A two-way ANOVA on ANDI revealed no significant main effects of condition, $F(1, 27) = 0.19$, $p = .665$, $f = .08$, or precision, $F(1, 27) = 1.70$, $p = .203$, $f = .14$, and no significant interaction, $F(1, 27) = 0.03$, $p = .869$, $f = .09$. Mean ANDI for precise responses was .33 (SD = .40) in the eyes-open condition and .37 (SD = .41) in the eyes-closed condition, and for imprecise responses was .34 (SD = .47) in the eyes-open condition and .23 (SD = .40) in the eyes-closed condition.

2.3. Discussion

Increased retention intervals and repeated recall attempts both contribute to confidence inflation (e.g., Sauer et al., 2010; Shaw & McClure, 1996). Nevertheless, even under these conditions, eye-closure improved accuracy without significantly affecting mean confidence or discrimination between accurate and inaccurate responses. Thus, we replicated the findings from Experiment 1 with different stimulus materials, a longer retention interval, and repeated recall attempts.

3. Experiment 3

In Experiment 3, we analyzed the confidence ratings collected by Vredeveldt et al. (2011), who investigated how cued recall of a violent video-taped event was affected by environmental distractions during the interview. They compared four interview conditions: closing the eyes (no visual distraction), looking at a blank screen (minimal visual distraction), being exposed to visual stimuli (high visual distraction), and being exposed to auditory stimuli (high auditory distraction). They found both a general effect of distraction on recall performance (i.e., visual or auditory distraction reduced performance compared to closing the eyes or looking at a blank screen) and a modality-specific effect (i.e., recall of visual information was most disrupted by visual distraction, whereas recall of auditory information was most disrupted by auditory distraction). The non-significant difference in recall performance between closing the eyes and looking at a blank screen suggested that the eye-closure effect is due to reduced distraction in the environment, rather than the act of closing the eyes itself. This lends credence to Fisher and Geiselman's (1992) recommendation that, in a witness interview setting, looking at a blank wall (instead of at the interviewer) may be a suitable alternative to eye-closure.

The five-point confidence scales used in Experiments 1 and 2 did not permit calibration analysis. In Experiment 3, we measured confidence on the same scale as accuracy (i.e., 0–100%), which allowed us to analyze the CA relation using the calibration approach. This approach involves plotting the subjective probability of a response being correct (i.e., confidence) against the objective probability of a response being correct (i.e., accuracy). Thus, perfect calibration is achieved if 80% of responses provided with 80% confidence are accurate, 50% of responses provided with 50% confidence are accurate, and so on. A visual comparison with this ideal permits an assessment of the realism and linearity of the obtained function. Additionally, the calibration approach includes statistical indices of the CA relation (see Baranski & Petrusic, 1994; Juslin et al., 1996). The calibration (C) statistic reflects how close a witness is to perfect calibration, ranging from 0 (perfect calibration) to 1 (no calibration), and the over/underconfidence (O/U) statistic assesses the degree to which individuals are generally less or more confident than they are accurate, ranging from -1 (underconfidence) to 1 (overconfidence). We examined the effects of eye-closure on the realism of participants' confidence assessments, and on their general tendency to display overconfidence (e.g., Gigerenzer et al., 1991).

3.1. Method

Vredeveldt et al. (2011) provide a detailed description of the methodology. In brief, 80 native English speakers watched a video depicting a violent encounter between survivors on an island, completed a five-minute distracter task, and participated in a cued-recall interview with 20 questions about the event. After each

Please cite this article in press as: Vredeveldt, A., & Sauer, J.D. Effects of eye-closure on confidence-accuracy relations in eyewitness testimony. *Journal of Applied Research in Memory and Cognition* (2015), http://dx.doi.org/10.1016/j.jarmac.2014.12.006

Table 1
Mean values for the adjusted normalized discrimination index (ANDI), calibration statistic (C), and over-/underconfidence (O/U) statistic for precise and imprecise responses as a function of interview condition in Experiment 3.

Measure	**Condition**			
	Eyes closed	Blank screen	Visual distraction	Auditory distraction
ANDI, precise	.48 (.44)	.55 (.34)	.41 (.35)	.37 (.37)
ANDI, imprecise	.48 (.49)	.36 (.48)	.80 (.36)	.54 (.45)
C, precise	.05 (.03)	.07 (.05)	.07 (.04)	.06 (.04)
C, imprecise	.19 (.20)	.13 (.09)	.16 (.11)	.09 (.06)
O/U, precise	−.04 (.09)	−.03 (.12)	.07 (.11)	.08 (.13)
O/U, imprecise	−.25 (.26)	−.21 (.17)	−.17 (.20)	−.09 (.18)

Note. Standard deviations in parentheses.

response, participants indicated their confidence on a scale of 0% ("not confident at all") to 100% ("extremely confident"). Participants were randomly assigned to one of four interview conditions, in which they were instructed to (a) keep their eyes closed, (b) look at a blank computer screen, (c) look at a computer screen on which Hebrew words appeared at a rate of one per second, or (d) look at a blank computer screen while hearing the same Hebrew words being spoken via speakers. For the present research, audiotaped interviews were re-coded using the procedure described for Experiment 1. Sixteen interviews (320 responses; 20% of the total sample) were randomly selected and scored independently by a second blind coder. Interrater reliability was high, $\kappa = .86$, $p < .001$. The codes of the first coder were retained.

3.2. Results

3.2.1. Recall accuracy

A 4 (Interview Condition: eyes closed, blank screen, visual distraction, auditory distraction) by 2 (Precision: precise, imprecise) mixed ANOVA on proportion correct revealed that imprecise answers ($M = .87$, SD = .15) were again significantly more accurate than precise answers ($M = .81$, SD = .12), F (1, 75) = 24.09, $p < .001$, $f = .39$. There was also a significant main effect of condition, F (3, 75) = 3.59, $p = .018$, $f = .38$, but no significant interaction, F (3, 75) = 0.41, $p = .746$, $f = .09$. Three simple ANOVAs (Bonferroni-corrected $\alpha = .017$) revealed no significant difference in accuracy between the eyes-closed ($M = .88$, SD = .08) and blank-screen ($M = .84$, SD = .08) conditions, F (1, 38) = 3.58, $p = .066$, $f = .31$, but participants who closed their eyes were significantly more accurate compared to participants in the visual-distraction ($M = .79$, SD = .09), F (1, 38) = 12.40, $p = .001$, $f = .57$, and auditory-distraction ($M = .78$, SD = .12) conditions, F (1, 38) = 10.38, $p = .003$, $f = .52$. Interview condition did not significantly affect the number of "don't know" responses, F (1, 38) = 0.63, $p = .599$, $f = .16$.

3.2.2. Mean confidence

As in Experiments 1 and 2, multilevel Model C best fit the mean confidence data, reducing between-subjects variability by 22.8% and within-subjects variability by 5.2%. Participants were again significantly more confident in accurate ($M = 84.29$, SD = 19.96) than inaccurate responses ($M = 60.81$, SD = 29.06), F (1, 1443.78) = 300.47, $p < .001$, $f = .46$, and significantly more confident in precise ($M = 83.64$, SD = 21.34) than imprecise responses ($M = 70.60$, SD = 26.69), F (1, 1448.26) = 147.47, $p < .001$, $f = .32$. This time, there was also a significant main effect of interview condition on mean confidence, F (1, 79.84) = 5.00, $p = .028$, $f = .25$. Nevertheless, three separate multilevel models (Bonferroni-corrected $\alpha = .017$) revealed no significant differences in mean confidence between the eyes-closed condition ($M = 81.69$, SD = 20.95) and the blank-screen condition ($M = 77.73$, SD = 26.07), F (1, 39.89) = 0.08, $p = .778$, $f = .05$, visual-distraction condition ($M = 78.73$, SD = 24.84), F (1, 39.67) = 0.30, $p = .585$, $f = .09$, or auditory-distraction condition ($M = 82.43$, SD = 21.92), F (1, 39.95) = 4.40, $p = .042$, $f = .33$, respectively. Although none of the individual comparisons involving eye-closure showed a significant difference at the Bonferroni-corrected α level, it is worth noting that the difference between the eyes-closed and auditory-distraction conditions was marginally significant and moderate in size. The mean confidence ratings look similar, but participants in the auditory-distraction condition displayed relatively high confidence in light of their lower levels of accuracy and precision.

3.2.3. Discrimination between accurate and inaccurate responses

A two-way ANOVA on ANDI revealed no significant main effects of condition, F (3, 34) = 1.52, $p = .227$, $f = .37$, or precision, F (1, 34) = 0.78, $p = .385$, $f = .11$, and no significant interaction, F (3, 34) = 2.28, $p = .097$, $f = .30$ (see Table 1).

3.2.4. Confidence-accuracy calibration

Fig. 1 shows CA calibration curves for precise responses and Fig. 2 for imprecise responses. Confidence ratings were collapsed into eleven categories (0–9%, 10–19%, . . ., 90–99%, and 100%). Percentage correct in each category was compared with the weighted mean confidence rating in that category. A two-way ANOVA on the C statistic revealed a significant effect of precision, F (1, 76) = 37.83, $p < .001$, $f = .46$, but no significant effect of interview condition, F (3, 76) = 2.01, $p = .120$, $f = .28$, and no significant interaction, F (3, 76) = 2.40, $p = .074$, $f = .20$. Table 1 shows that confidence and accuracy were much better calibrated (i.e., closer to 0) for precise responses than for imprecise responses. Indeed, precise responses (Fig. 1) produce generally linear, positive CA relations (regardless of interview condition), whereas for imprecise responses (Fig. 2) no meaningful CA relationship is evident.

A two-way ANOVA on O/U revealed significant effects of interview condition, F (3, 76) = 5.02, $p = .003$, $f = .45$, and precision, F (1, 76) = 79.54, $p < .001$, $f = .62$, but no significant interaction, F (3, 76) = 0.41, $p = .744$, $f = .08$. Underconfidence was greater for imprecise than precise responses. Three simple ANOVAs (Bonferroni-corrected $\alpha = .017$) revealed no significant difference between the eyes-closed condition and the blank-screen condition, F (1, 38) = 0.00 $p = .988$, $f = .00$, but a marginally significant difference between eye-closure and visual distraction, F (1, 38) = 4.35, $p = .044$, $f = .34$, and a significant difference between eye-closure and auditory distraction, F (1, 38) = 11.58, $p = .002$, $f = .55$. Distraction was generally associated with increased overconfidence (see Table 1 and Fig. 1).

3.3. Discussion

Consistent with Experiments 1 and 2, eye-closure (cf. visual- and auditory-distraction) significantly improved recall accuracy. There were no significant differences between eye-closure and the other interview conditions in terms of mean confidence, discrimination between accurate and inaccurate responses, or CA calibration.

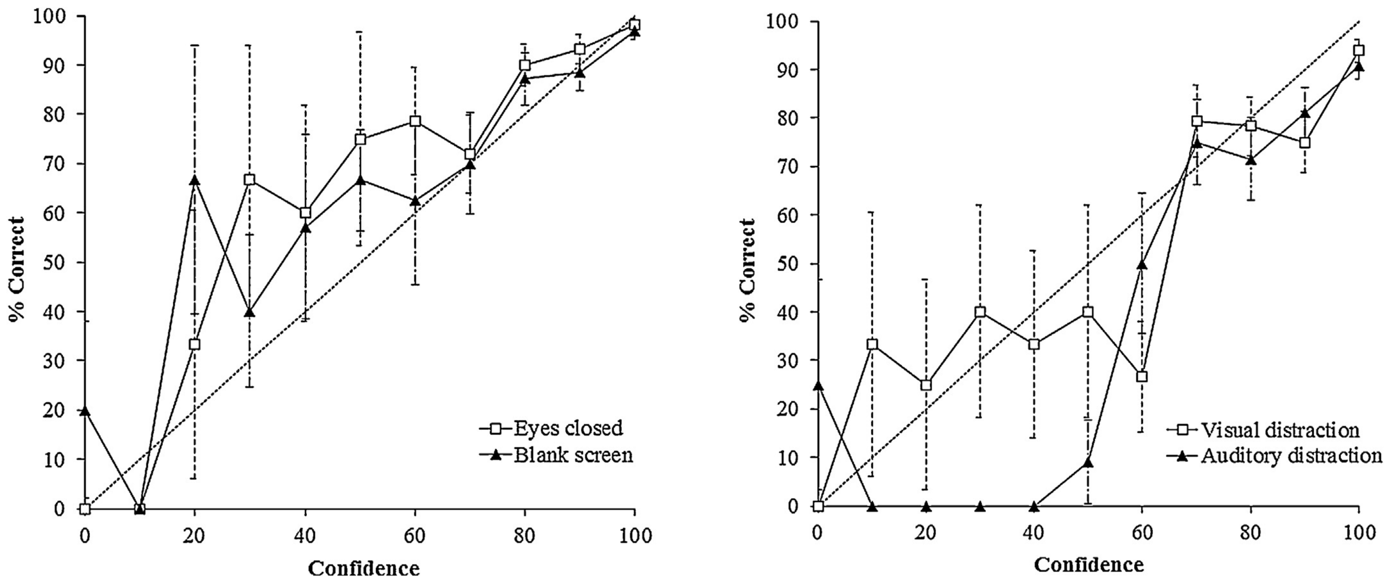

Fig. 1. Confidence-accuracy calibration for precise responses in Experiment 3, showing percentage correct as a function of confidence rating (both measured on a scale of 0–100%). Left panel shows eyes-closed (EC) and blank-screen (BS) conditions, and right panel shows visual-distraction (VD) and auditory-distraction (AD) conditions. The dotted black line denotes perfect calibration. Error bars indicate standard error.

However, reducing distraction significantly decreased overconfidence.

4. General discussion

Three experiments investigated the effect of eye-closure on the confidence-accuracy relation for witnesses' event memory. The findings were remarkably consistent across experiments. First, participants were able to monitor the accuracy of their responses, indicated by both the mean ANDI statistics and the higher confidence ratings for correct than incorrect responses. Second, participants were significantly less confident in imprecise responses than in precise responses, and showed poorer calibration for imprecise responses. Third, across all experiments, eye-closure reliably increased recall accuracy without significantly inflating confidence, consistent with Wagstaff et al. (2004), Wagstaff, Wheatcroft, Burt, et al., (2011) and Wagstaff, Wheatcroft, and Caddick (2011) findings. The present research included discrimination and calibration analyses as additional indices of the CA relation. Although eye-closure did not improve participants' discrimination between correct and incorrect responses, there was no evidence that it impaired this ability. Further, reducing distraction in Experiment 3 significantly reduced overconfidence. In sum, eye-closure improves recall accuracy with no apparent cost, and some evidence of benefit, to metamemory.

We can make some observations regarding potential theoretical underpinnings of the observed effects. Confidence assessments can be influenced by intrinsic cues, which are tied to the to-be-remembered stimuli, and extrinsic cues, which relate to the external environment associated with encoding and retrieval (Koriat, 1995, 1997). Whereas intrinsic cues, such as ease of retrieval and decision time, typically affect confidence and accuracy in equivalent ways (Sauerland & Sporer, 2009; see also Kelley & Lindsay, 1993), extrinsic cues often exert disproportionate effects

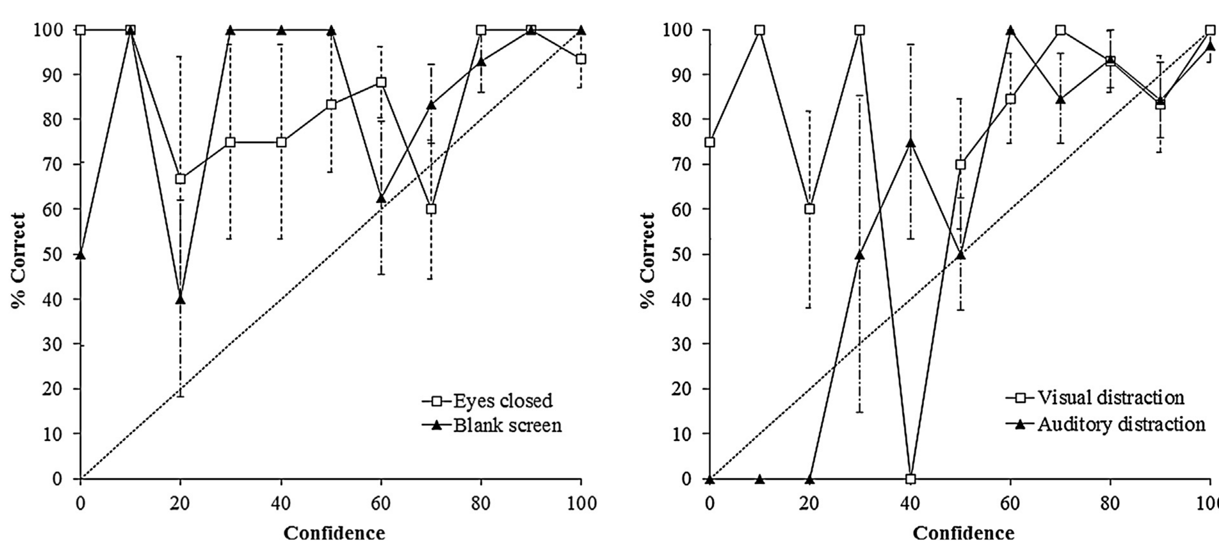

Fig. 2. Confidence-accuracy calibration for imprecise responses in Experiment 3, showing percentage correct as a function of confidence rating (both measured on a scale of 0–100%). Left panel shows eyes-closed (EC) and blank-screen (BS) conditions, and right panel shows visual-distraction (VD) and auditory-distraction (AD) conditions. The dotted black line denotes perfect calibration. Error bars indicate standard error.

on confidence (Busey, Tunnicliff, Loftus, & Loftus, 2000; Chandler, 1994; Garrioch & Brimacombe, 2001; Tulving, 1981). It is possible that eye-closure shifts witnesses' focus toward internal mental processes. For example, eye-closure polarizes moral judgments (Caruso & Gino, 2011) and intensifies emotional responses to negative emotional music (Lerner, Papo, Zhdanov, Belozersky, & Hendler, 2009). An enhanced focus on intrinsic cues as a result of eye-closure could explain the observed increase in accuracy (Koriat, 1993, 1995, 1997), as well as the decrease in overconfidence (cf. Arkes, Christensen, Lai, & Blumer, 1987). Moreover, even if eye-closure does not direct attention toward diagnostic internal cues, it should at least mitigate the effects of any non-diagnostic visual cues (e.g., from an interviewer). Other manipulations that encourage reflection on internal memory processes, such as focused meditation, have similarly been found to increase recall accuracy without inflating confidence (e.g., Hammond et al., 2006; Wagstaff et al., 2004; Wagstaff, Wheatcroft, & Caddick, 2011), improve CA relations (Brewer, Keast, & Rishworth, 2002; Kassin, 1985; Kassin, Rigby, & Castillo, 1991), and reduce overconfidence (Arkes et al., 1987; Buratti & Allwood, 2012a,b).

It is unclear, however, why other procedures that involve an eye-closure instruction do not have the same effect. Hypnosis, mental context reinstatement, and, to an extent, the CI, all increase overconfidence (Granhag et al., 2004; Hammond et al., 2006; Kebbell & Wagstaff, 1998). Perhaps these complex procedures raise witnesses' expectations with regard to the quality of their memory reports (e.g., "this procedure will help me to remember accurate information") to a greater extent than the eye-closure instruction on its own. External cues provided by the interviewer may inflate confidence without affecting accuracy (for more on expectancy effects, see Lynn & Nash, 1994; Wagstaff, 2008; Wagstaff, Vella, & Perfect, 1992; Webert, 2003). Participants in the present experiments were not informed *why* they were instructed to close their eyes, in contrast with, for example, the eye-closure instruction in the Cognitive Interview manual ("you'll probably find it easier to concentrate if you close your eyes"; Fisher & Geiselman, 1992, p. 133).

4.1. Practical application

The finding that participants were able to monitor the accuracy of their responses suggests that the use of eyewitness confidence to inform investigative and legal decisions may not be as inappropriate as once thought (see e.g., Bothwell, Deffenbacher, & Brigham, 1987; Penrod, Loftus, & Winkler, 1982; Wells & Murray, 1984). Of course, in legal settings, confidence must be considered alongside other cues to accuracy, such as witnessing conditions (e.g., exposure duration, distance and lighting) and external influences (e.g., retention interval, feedback from others). Precise responses were associated with greater overconfidence (or less underconfidence) than imprecise responses (see also Goldsmith et al., 2002; 2005; Weber & Brewer, 2008). Accordingly, legal professionals may need to regard highly specific details provided with high confidence with more caution than relatively imprecise details provided with high confidence. Further, eye-closure improved recall accuracy of precise details but not of imprecise details (see also Vredeveldt & Penrod, 2013). Thus, the most important application of the eye-closure instruction may be to help witnesses remember specific details about an event.

The key practical implication of the present research is that reducing distraction during recall (through eye-closure) improves accuracy without significantly inflating confidence or impairing CA relations. Given that police investigators, legal professionals, and jurors give considerable weight to witness confidence, a recall enhancement procedure that maintains the validity of confidence judgments is desirable. The eye-closure instruction is easy to implement in practice (Vredeveldt, Tredoux, et al., 2014), does not require training resources, and does not extend interview time. Moreover, where hypnosis, mental context reinstatement, and, to an extent, the Cognitive Interview, all increase overconfidence, we found no evidence that eye-closure inflates confidence. On the contrary, reducing distractions in Experiment 3 significantly *reduced* overconfidence.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Author note

Part of this research was supported by a University of York departmental studentship and a VSB foundation scholarship awarded to Annelies Vredeveldt. We thank Alan Baddeley, Graham Hitch, and Ron Fisher for their valuable feedback on earlier versions of this work, and Jodie Davies-Thompson, Yvonna Lavis, Alieke Hildebrandt, and Sabrina de Haseth for their help with coding the data.

References

Allwood, C. M., Ask, K., & Granhag, P. A. (2005). The Cognitive Interview: Effects on the realism in witnesses' confidence in their free recall. *Psychology, Crime & Law*, *11*, 183–198. http://dx.doi.org/10.1080/10683160512331329943

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, *39*, 133–144. http://dx.doi.org/10.1016/0749-5978(87)90049-5

Baranski, J., & Petrusic, W. (1994). The calibration and resolution of confidence in perceptual judgments. *Attention, Perception, & Psychophysics*, *55*, 412–428. http://dx.doi.org/10.3758/bf03205299

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. http://dx.doi.org/10.1037/0021-9010.72.4.691

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior*, *26*, 353–364. http://dx.doi.org/10.1023/a:1015380522722

Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied*, *8*, 44–56. http://dx.doi.org/10.1037/1076-898X.8.1.44

Brigham, J. C., & Wolfskeil, M. P. (1983). Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identification. *Law and Human Behavior*, *7*, 337–349. http://dx.doi.org/10.1007/BF01044736

Buratti, S., & Allwood, C. M. (2012a). The accuracy of meta-metacognitive judgments: Regulating the realism of confidence. *Cognitive Processing*, *13*, 243–253. http://dx.doi.org/10.1007/s10339-012-0440-5

Buratti, S., & Allwood, C. M. (2012b). Improved realism of confidence for an episodic memory event. *Judgment & Decision Making*, *7*, 590–601.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*, 26–48. http://dx.doi.org/10.3758/bf03210724

Caruso, E. M., & Gino, F. (2011). Blind ethics: Closing one's eyes polarizes moral judgments and discourages dishonest behavior. *Cognition*, *118*, 280–285. http://dx.doi.org/10.1016/j.cognition.2010.11.008

Chandler, C. C. (1994). Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition test. *Memory & Cognition*, *22*, 273–280. http://dx.doi.org/10.3758/bf03200854

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, *14*, 185–191. http://dx.doi.org/10.1007/bf01062972

Dywan, J., & Bowers, K. (1983). The use of hypnosis to enhance recall. *Science*, *222*, 184–185. http://dx.doi.org/10.1126/science.6623071

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 552–564. http://dx.doi.org/10.1037/0096-1523.3.4.552

Fisher, R. P. (1996). Implications of output-bound measures for laboratory and field research in memory. *Behavioral and Brain Sciences*, *19*, 197–197. http://dx.doi.org/10.1017/S0140525X00042229

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield: Charles Thomas.

Garrioch, L., & Brimacombe, C. (2001). Lineup administrators' expectations: Their impact on eyewitness confidence. *Law and Human Behavior*, *25*, 299–315. http://dx.doi.org/10.1023/a:1010750028643

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528. http://dx.doi.org/10.1037/0033-295x.98.4.506

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language*, *52*, 505–525. http://dx.doi.org/10.1016/j.jml.2005.01.010

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). Strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General*, *131*, 73–95. http://dx.doi.org/10.1037//0096-3445.131.1.73

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, *49*, 732–764. http://dx.doi.org/10.1080/01621459.1954.10501231

Granhag, P. A. (1997). Realism in eyewitness confidence as a function of type of event witnessed and repeated recall. *Journal of Applied Psychology*, *82*, 599–613. http://dx.doi.org/10.1037/0021-9010.82.4.599

Granhag, P. A., Jonsson, A.-C., & Allwood, C. M. (2004). The cognitive interview and its effect on witnesses' confidence. *Psychology, Crime & Law*, *10*, 37–52. http://dx.doi.org/10.1080/10683160210000030577

Gwyer, P., & Clifford, B. R. (1997). The effects of the cognitive interview on recall, identification, confidence and the confidence/accuracy relationship. *Applied Cognitive Psychology*, *11*, 121–145. http://dx.doi.org/10.1002/(sici)1099-0720(199704)11:2<121::aid-acp443>3.0.co;2-l

Hammond, L., Wagstaff, G. F., & Cole, J. (2006). Facilitating eyewitness memory in adults and children with context reinstatement and focused meditation. *Journal of Investigative Psychology and Offender Profiling*, *3*, 117–130. http://dx.doi.org/10.1002/jip.47

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1304–1316. http://dx.doi.org/10.1037/0278-7393.22.5.1304

Kassin, S. M. (1985). Eyewitness identification: Retrospective self-awareness and the accuracy–confidence correlation. *Journal of Personality and Social Psychology*, *49*, 878–893. http://dx.doi.org/10.1037/0022-3514.49.4.878

Kassin, S. M., Rigby, S., & Castillo, S. R. (1991). The accuracy-confidence correlation in eyewitness testimony: Limits and extensions of the retrospective self-awareness effect. *Journal of Personality and Social Psychology*, *61*, 698–707. http://dx.doi.org/10.1037/0022-3514.61.5.698

Kebbell, M. R., & Wagstaff, G. F. (1998). Hypnotic interviewing: The best way to interview eyewitnesses? *Behavioral Sciences & the Law*, *16*, 115–129. http://dx.doi.org/10.1002/(SICI)1099-0798(199824)16:1<115::AID-BSL296>3.0.CO;2-I

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, *32*, 1–24. http://dx.doi.org/10.1006/jmla.1993.1001

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639. http://dx.doi.org/10.1037/0033-295x.100.4.609

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*, 311–333. http://dx.doi.org/10.1037/0096-3445.124.3.311

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370. http://dx.doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*, 80–113. http://dx.doi.org/10.1037/a0025648

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517. http://dx.doi.org/10.1037/0033-295X.103.3.490

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 107–118. http://dx.doi.org/10.1037/0278-7393.6.2.107

Lerner, Y., Papo, D., Zhdanov, A., Belozersky, L., & Hendler, T. (2009). Eyes wide shut, Amygdala mediates eyes-closed effect on emotional experience with music. *PLoS ONE*, *4*, e6230. http://dx.doi.org/10.1371/journal.pone.0006230

Lynn, S. J., & Nash, M. R. (1994). Truth in memory: Ramifications for psychotherapy and hypnotherapy. *American Journal of Clinical Hypnosis*, *36*, 194–208. http://dx.doi.org/10.1080/00029157.1994.10403069

Masson, M. J. E., & Rotello, C. M. (2009). *Sources of bias in the Goodman–Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes*, *35*, 509–527. http://dx.doi.org/10.1037/a0014876

McCauley, M. R., & Fisher, R. P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology*, *80*, 510–516. http://dx.doi.org/10.1037/0021-9010.80.4.510

McMahon, M. (2000). The effect of the enhanced cognitive interview on recall and confidence in elderly adults. *Psychiatry, Psychology and Law*, *7*, 9–32. http://dx.doi.org/10.1080/13218710009524968

Mello, E. W., & Fisher, R. P. (1996). Enhancing older adult eyewitness memory with the cognitive interview. *Applied Cognitive Psychology*, *10*, 403–417. http://dx.doi.org/10.1002/(SICI)1099-0720(199610)10:5<403::AID-ACP395>3.0.CO;2-X

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259. http://dx.doi.org/10.1037/0033-295x.84.3.231

Noon, E., & Hollin, C. R. (1987). Lay knowledge of eyewitness behaviour: A British survey. *Applied Cognitive Psychology*, *1*, 143–153. http://dx.doi.org/10.1002/acp.2350010207

Odinot, G., & Wolters, G. (2006). Repeated recall, retention interval and the accuracy–confidence relation in eyewitness memory. *Applied Cognitive Psychology*, *20*, 973–985. http://dx.doi.org/10.1002/acp.1263

Odinot, G., Wolters, G., & Lavender, T. (2009). Repeated partial eyewitness questioning causes confidence inflation but not retrieval-induced forgetting. *Applied Cognitive Psychology*, *23*, 90–97. http://dx.doi.org/10.1002/acp.1443

Odinot, G., Wolters, G., & van Giezen, A. (2013). Accuracy, confidence and consistency in repeated recall of events. *Psychology, Crime and Law*, *19*, 629–642. http://dx.doi.org/10.1080/1068316X.2012.660152

Penrod, S. D., Loftus, E. F., & Winkler, J. (1982). The reliability of eyewitness testimony: A psychological perspective. In N. L. Kerr, & R. M. Bray (Eds.), *The psychology of the courtroom* (pp. 119–168). New York: Academic Press.

Perfect, T. J., Wagstaff, G. F., Moore, D., Andrews, B., Cleveland, V., Newcombe, S., et al. (2008). How can we help witnesses to remember more? It's an (eyes) open and shut case. *Law and Human Behavior*, *32*, 314–324. http://dx.doi.org/10.1007/s10979-007-9109-5

Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement*, *65*, 717–741. http://dx.doi.org/10.1177/0013164405278558

Roberts, W. T., & Higham, P. A. (2002). Selecting accurate statements from the cognitive interview using confidence ratings. *Journal of Experimental Psychology: Applied*, *8*, 33–43. http://dx.doi.org/10.1037/1076-898X.8.1.33

Sauer, J. D., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, *34*, 337–347. http://dx.doi.org/10.1007/s10979-009-9192-x

Sauerland, M., & Sporer, S. L. (2009). Fast and confident: Postdicting eyewitness identification accuracy in a field study. *Journal of Experimental Psychology: Applied*, *15*, 46–62. http://dx.doi.org/10.1037/a0014560

Shaw, J. S., III. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, *2*, 126–146. http://dx.doi.org/10.1037/1076-898x.2.2.126

Shaw, J. S., III, & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior*, *20*, 629–653. http://dx.doi.org/10.1007/bf01499235

Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior*, *20*, 479–496. http://dx.doi.org/10.1016/S0022-5371(81)90129-8

Vredeveldt, A., Baddeley, A. D., & Hitch, G. J. (2014). The effectiveness of eye-closure in repeated interviews. *Legal and Criminological Psychology*, *19*, 282–295. http://dx.doi.org/10.1111/lcrp.12013

Vredeveldt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eyeclosure helps memory by reducing cognitive load and enhancing visualisation. *Memory & Cognition*, *39*, 1253–1263. http://dx.doi.org/10.3758/s13421-011-0098-8

Vredeveldt, A., & Penrod, S. D. (2013). Eye-closure improves memory for a witnessed event under naturalistic conditions. *Psychology, Crime & Law*, *19*, 893–905. http://dx.doi.org/10.1080/1068316x.2012.700313

Vredeveldt, A., Tredoux, C. G., Nortje, A., Kempen, K., Puljević, C., & Labuschagne, G. N. (2014). A field evaluation of the Eye-Closure Interview with witnesses of serious crimes. *Law and Human Behavior*, http://dx.doi.org/10.1037/lhb0000113 (Advance online publication)

Wagstaff, G. F. (2008). Hypnosis and the law: Examining the stereotypes. *Criminal Justice and Behavior*, *35*, 1277–1294. http://dx.doi.org/10.1177/0093854808321669

Wagstaff, G. F., Brunas-Wagstaff, J., Cole, J., Knapton, L., Winterbottom, J., Crean, V., et al. (2004). Facilitating memory with hypnosis, focused meditation, and eye closure. *International Journal of Clinical & Experimental Hypnosis*, *52*, 434–455. http://dx.doi.org/10.1080/00207140490889062

Wagstaff, G. F., Vella, M., & Perfect, T. (1992). The effect of hypnotically elicited testimony on jurors' judgments of guilt and innocence. *The Journal of Social Psychology*, *132*, 591–595. http://dx.doi.org/10.1080/00224545.1992.9713898

Wagstaff, G. F., Wheatcroft, J. M., Burt, C. L., Pilkington, H. J., Wilkinson, K., & Hoyle, J. D. (2011). Enhancing witness memory with focused meditation and eye-closure: Assessing the effects of misinformation. *Journal of Police and Criminal Psychology*, *26*, 152–161. http://dx.doi.org/10.1007/s11896-010-9082-7

Wagstaff, G. F., Wheatcroft, J. M., Caddick, A. M., Kirby, L. J., & Lamont, E. (2011). Enhancing witness memory with techniques derived from hypnotic investigative interviewing: Focused meditation, eye-closure, and context reinstatement. *International Journal of Clinical and Experimental Hypnosis*, *59*, 146–164. http://dx.doi.org/10.1080/00207144.2011.546180

Weber, N., & Brewer, N. (2008). Eyewitness recall: Regulation of grain size and the role of confidence. *Journal of Experimental Psychology: Applied*, *14*, 50–60. http://dx.doi.org/10.1037/1076-898X.14.1.50

Webert, D. R. (2003). Are the courts in a trance? Approaches to the admissibility of hypnotically enhanced witness testimony in the light of empirical evidence. *American Criminal Law Review*, *40*, 1301–1327.

Wells, G. L., & Murray, D. M. (1984). Eyewitness confidence. In G. L. Wells, & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 155–170). New York: Cambridge University Press.

Yaniv, I., Yates, J. F., & Smith, J. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin*, *110*, 611–617. http://dx.doi.org/10.1037/0033-2909.110.3.611

Please cite this article in press as: Vredeveldt, A., & Sauer, J.D. Effects of eye-closure on confidence-accuracy relations in eyewitness testimony. *Journal of Applied Research in Memory and Cognition* (2015), http://dx.doi.org/10.1016/j.jarmac.2014.12.006