Journal of Applied Research in Memory and Cognition 4 (2015) 51–58

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 6 January 2015

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

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).