


## RESEARCH ARTICLE

WILEY

# Improving the evaluation of eyewitness evidence in legal decision-making: Testing an active versus passive teaching aid

Renan Saraiva<sup>1</sup>  | Giulia Bertoldo<sup>2</sup> | Ludvig Daae Bjørndal<sup>3</sup> | Cătălina Bunghez<sup>4</sup> | Ingvild Sandø Lofthus<sup>3</sup> | Lucy McGill<sup>5</sup> | Stéphanie Richardson<sup>6</sup> | Marie Stadel<sup>7</sup>

<sup>1</sup>Department of Psychology, University of Portsmouth, Portsmouth, UK

<sup>2</sup>Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova, Padua, Italy

<sup>3</sup>Department of Psychology, University of Oslo, Oslo, Norway

<sup>4</sup>Department of Psychology, West University of Timisoara, Timisoara, Romania

<sup>5</sup>Department of Psychology, Trinity College Dublin, Dublin, Ireland

<sup>6</sup>School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK

<sup>7</sup>Department of Sociology, University of Groningen, Groningen, The Netherlands

## Correspondence

Renan Saraiva, Department of Psychology, University of Portsmouth, Portsmouth, UK. Email: [renan.saraiva@port.ac.uk](mailto:renan.saraiva@port.ac.uk)

## Funding information

Association for Psychological Science

## Abstract

Judges, jurors and other triers of fact often rely upon eyewitness evidence in criminal trials, but eyewitness memory is not always accurate and can sometimes be contaminated. The I-I-Eye is an evidence-based teaching aid designed to improve the evaluation of eyewitness evidence in legal settings. We aimed to further test the I-I-Eye and examine whether adding an active component to this teaching aid improves its effectiveness. Two experiments ( $N = 324$  and  $N = 322$ ) were conducted using a 2 (case strength: weak vs. strong) by 3 (teaching aid condition: control vs. passive vs. active) between-subjects design. Results of both experiments showed that the I-I-Eye can help jurors recognize strong eyewitness cases, although it was not particularly effective when the evidence was weak. It was also found that the active component did not further improve sensitivity. We discuss whether teaching aids such as the I-I-Eye may assist decision-makers in the evaluation of eyewitness evidence, while highlighting some of its main limitations found in our results.

## KEYWORDS

eyewitness memory, eyewitness testimony, I-I-Eye, jury decision-making, teaching aids

## 1 | INTRODUCTION

Eyewitness testimony plays a major role in the criminal justice system and is often an influential type of evidence during trials. Judges and juries commonly rely on eyewitness accounts to make decisions regarding the guilt and sentencing of a defendant. However, studies have shown that triers of fact are often unaware of factors that can impair eyewitness memory accuracy (Bjørndal et al., 2020; Bradfield et al., 2002; Magnussen et al., 2010). This is of great concern as inaccurate eyewitness statements or identifications could lead to wrongful convictions and miscarriages of justice (American Psychological Association, 2012; Wells et al., 1998). As a result, eyewitness memory researchers have investigated the use of judicial instructions and teaching aids, aiming to improve the evaluation of eyewitness evidence by

triers of fact. Much of that research has found that most judicial instructions do not improve jurors ability to discern eyewitness evidence quality; rather, jurors receiving those instructions tend to be skeptical of all eyewitness evidence (e.g., Jones et al., 2017; Jones et al., 2020). However, one specific teaching aid that has shown promising results in sensitizing jurors to weak and strong eyewitness evidence is the I-I-Eye (which stands for interview, identification, and eyewitness factors; Pawlenko et al., 2013; Safer et al., 2016). However, current evidence on the I-I-Eye's effectiveness is still sparse and potential improvements to the I-I-Eye which may further assist jurors in applying their knowledge to criminal trials have yet to be tested. These improvements include, for example, requiring trainees to actively engage with the learning materials. In the current study, we aimed to collect and examine new evidence on the effectiveness of the I-I-Eye,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Applied Cognitive Psychology* published by John Wiley & Sons Ltd.

while also testing whether an active component would improve this teaching aid.

## 1.1 | Issues with eyewitness memory

It has been established for many years now that eyewitness recollections are not a permanent record of perceived events and may be tainted by a number of different factors. That is, witnesses' memories of important details not only decline over time (Sauer et al., 2010), but can also be negatively distorted by new information introduced after the original experience. The use of leading questions during criminal investigations, for example, may influence witnesses to disclose information that they never actually encountered (Roebbers & Schneider, 2000). Even during the crime itself, many factors such as crime duration (Memon et al., 2003), distance (Nyman et al., 2019), and the presence of a weapon (Fawcett et al., 2013) can negatively impact how witnesses will remember the event. Factors such as these can contribute to faulty witness testimony and misidentifications of innocent suspects. Such an issue has serious implications as some records suggest that misidentifications have played a role in over 70% of known exonerations based on DNA evidence in the USA alone (Innocence Project, 2021).

However, that is not to say that eyewitness evidence is inherently unreliable. As with other types of evidence (e.g., fingerprints or DNA), eyewitness evidence may or may not be contaminated, depending on factors related to: (a) the interview, (b) the crime itself, and (c) the identification procedures (Wixted et al., 2018). In fact, uncontaminated eyewitness evidence can be crucial in solving crimes and prosecuting perpetrators, especially in cases lacking other physical evidence. Great measures have been taken to protect eyewitness evidence from contamination, including worldwide policy reforms (National Research Council, 2014; Police and Criminal Evidence Act, 2017; Technical Working Group for Eyewitness Evidence, 2014), but in many criminal justice systems it is ultimately up to the jury or judge to evaluate eyewitness testimonies and make decisions based on that evidence. Even in jurisdictions where adequate interview and lineup procedures are encouraged, it is still possible that triers of fact are exposed to contaminated testimonies. It is therefore of great importance that jurors and judges are instructed to effectively discriminate between weak and strong eyewitness evidence.

## 1.2 | Evaluating eyewitness evidence

Jurors typically rely heavily on eyewitnesses' testimonies for their decisions (Desmarais & Read, 2011). However, numerous studies have shown that lay knowledge of eyewitness memory issues is limited in scope and highly divergent from experts' knowledge. Benton et al., (2006), for example, examined the knowledge of a juror representative sample on 30 statements about eyewitness issues, finding that this sample disagreed with the experts on 87% of the issues. In a similar type of study, it was found that juror experience, in terms of number of times serving as juror, did not correlate with eyewitness

knowledge (Magnussen et al., 2010), showing that simply serving on a criminal trial does not make a juror more sensitive to the fallibility of eyewitness testimony. But not all evidence points to such a high discrepancy between jurors and experts' knowledge. In a meta-analytic review of 23 surveys assessing lay knowledge of eyewitness issues, Desmarais and Read (2011) found that a majority of lay respondents agreed with experts in as many as 11 of the 16 items included in the review. Although in some cases jurors may demonstrate reasonable agreement with experts on certain eyewitness memory issues, there seems to be several factors for which lay knowledge is generally faulty. Furthermore, having appropriate knowledge alone may not be sufficient to safeguard against reliance on weak eyewitness evidence, given that lay persons may still find it difficult to apply knowledge regarding eyewitness issues to the case at hand (Desmarais & Read, 2011).

Evidently, gaps in jurors' knowledge regarding eyewitness memory issues can have harmful consequences, and therefore jurors may rely on weak eyewitness evidence when deciding someone's guilt. That is one reason why different legal safeguards have been put in practice to try and inform juries of issues related to eyewitness testimony, including expert testimony, judicial instructions and the main focus of the current research: teaching aids.

Teaching aids typically include information about factors that are known to affect eyewitness memory accuracy. Those may include elements that are related to: (a) the interview (e.g., leading questions; Roebbers & Schneider, 2000), (b) the identification task (e.g., biased lineup instructions; Clark, 2005), or (c) the crime (e.g., effect of weapon presence; Fawcett et al., 2013). Based on the current literature, one specific teaching aid that has been associated with promising results regarding the evaluation of eyewitness evidence is the I-I-Eye (Pawlenko et al., 2013; Safer et al., 2016). The original I-I-Eye (which stands for identification, interview, and eyewitness conditions) consists of a PowerPoint presentation (without audio or animations) in which triers of fact are taught to consider whether: (a) the eyewitness interview procedures were conducted properly; (b) the identification procedures were conducted properly; and (c) there were any factors present during the crime that might have negatively affected the eyewitness memory. Pawlenko et al., (2013) found that juries receiving this teaching aid had better sensitivity to both strong and weak eyewitness evidence, as demonstrated by fewer guilty verdicts in cases where the eyewitness memory was likely contaminated, compared to cases where the eyewitness memory was not contaminated.

## 1.3 | Issues related to judicial instructions and teaching aids

An important concern regarding teaching aids is that they might induce excessive skepticism among triers of fact, who consequently disregard all eyewitness evidence, even when there is reason to believe that the eyewitness account has not been contaminated (Dillon et al., 2017; Papailiou et al., 2015). This is partially because individuals may distrust any type of eyewitness evidence after learning of some factors that can make eyewitness memory unreliable. In

the United States, for example, a common jury instruction regarding eyewitness testimony is the Biggers factors, a list of five factors that jurors should consider when evaluating eyewitness memory accuracy (e.g., the eyewitness' view of the perpetrator during the crime). Research indicates that the Biggers factors instructions tend to make jurors skeptical of eyewitness testimony, rather than sensitize them to specific eyewitness factors in a case (Ramirez et al., 1996). A similar example can be observed for the Henderson Instructions, a set of judicial instructions proposed by The New Jersey Supreme Court to assist jurors in effectively evaluating eyewitness evidence (New Jersey v. Henderson, 2011). It has now been found in a number of different studies that instead of sensitizing jurors, the Henderson Instructions induce overall skepticism of eyewitness evidence (Dillon et al., 2017; Jones et al., 2017; Papailiou et al., 2015). However, it should be noted that Jones and Penrod (2018) have found that a modified version of the Henderson instructions can be effective in sensitizing jurors to eyewitness evidence, although their control instructions were roughly as effective in sensitizing jurors as the modified Henderson instructions.

The most promising aspect of the I-I-Eye as a teaching aid is that the current evidence suggests it does not create skepticism, as opposed to the aforementioned safeguards. Instead, it seems that the I-I-Eye increases sensitivity to both strong and weak eyewitness testimonies, although evidence for this finding is scarce (Pawlenko et al., 2013; Safer et al., 2016). Nevertheless, some improvements for the I-I-Eye have been proposed, including a phase in which evaluators actively apply the information they have learned to the case at hand (Wise et al., 2009). Jurors should not only have the knowledge of factors that influence eyewitness evidence, but should also be able to integrate such knowledge into their decision making. In educational research, active learning typically requires individuals to analyze, synthesize, and evaluate educational materials rather than passively encoding them, and have been shown to promote cognitive and meaningful engagement (Fredricks et al., 2004). Some theoretical frameworks predict better retention when individuals are actively, constructively or interactively engaged with the learning materials (e.g., ICAP framework; Chi & Wylie, 2014). The ICAP framework suggests that different modes of interacting with learning materials lead to different levels of learning outcomes, with Interactive modes of learning generating the best learning outcomes (e.g., class discussions), followed by Constructive (e.g., explaining to others or oneself), Active (e.g., summarizing the information) and Passive modes (e.g., reading text). In practice, further engagement can occur when the learning material is rehearsed, summarized, manipulated, or compared to prior knowledge and other materials. Some practical examples of Constructive behaviors include explaining to others or to oneself, taking notes in one's own words, posing problems, making predictions with the information learned, inducing hypotheses, self-evaluating or monitoring one's understanding, and so on (Chi et al., 2018). The ICAP framework has been used to improve cognitive engagement in teaching settings (Wiggins et al., 2017), education of health professionals (Lim et al., 2019), and video-based learning (Seo et al., 2021).

In the current study, it was theorized that the inclusion of an active form focused on applying the teaching aid information to a

specific case would promote further Active/Constructive learning modes, ultimately improving learning outcomes. Within the ICAP framework, active forms are assumed to include processes of activating prior knowledge, linking with activated prior knowledge, inferring from prior knowledge or the newly integrated knowledge, and storing the linked and inferred knowledge (Chi et al., 2018). However, previous studies have not specifically investigated the extent to which an active application component improves the evaluation of eyewitness accounts. The current study aimed to address this gap by testing whether individuals receiving an active component following the I-I-Eye teaching aid would be better able to discriminate between weak and strong eyewitness evidence.

## 1.4 | Current experiments

In the current study, we aimed to further test the effectiveness of the I-I-Eye and examine whether adding an active component to this teaching aid improves the evaluation of eyewitness testimony. We hypothesized that participants who receive the I-I-Eye teaching aid (either the active or passive version) would be able to better distinguish a weak eyewitness case from a strong eyewitness case than a control group who receives general judicial instructions (with no mention of eyewitness evidence). This will be reflected by fewer guilty verdicts in the teaching aid conditions for cases where the eyewitness evidence is weak as opposed to strong (H1). We expect participants who receive the active I-I-Eye aid to be more sensitive to the quality of eyewitness evidence, compared to those who receive the passive I-I-Eye aid. That is, we expect fewer guilty verdicts among those who received the active I-I-Eye teaching aid in cases with weak as opposed to strong eyewitness evidence, compared to those who received the passive I-I-Eye teaching aid (H2).

## 1.5 | Disclosures

For each experiment, we report how we determined our sample size, all data exclusions, all manipulations, and all materials. The studies in this manuscript were approved by the Faculty of Science Human Research Ethics Committee at the University of Portsmouth. Materials and data for Experiments 1 and 2 are available at: [https://osf.io/t46na/?view\\_only=ac69527fa8b04d10b9b374e1f01e3195](https://osf.io/t46na/?view_only=ac69527fa8b04d10b9b374e1f01e3195)

The preregistration is available at: [https://osf.io/b3fkv/?view\\_only=a62c88147384419da73c9c8285f44927](https://osf.io/b3fkv/?view_only=a62c88147384419da73c9c8285f44927)

## 2 | EXPERIMENT 1

### 2.1 | Method

#### 2.1.1 | Design

The experiment used a 3 (teaching aid: general judicial instruction vs. passive I-I-Eye aid vs. active I-I-Eye aid) × 2 (eyewitness evidence:

strong vs. weak case) between-subjects factorial design. Participants were randomly allocated to one of the six experimental conditions.

### 2.1.2 | Participants

The required sample size was estimated using a power analysis for logistic regressions (GPowerTM). Safer et al., (2016) found that I-I-Eye participants were 3.42 times more likely to render a guilty verdict in the strong case than the weak case compared to a control group who received general judicial instructions. We used this finding as a proxy for our target effect size, but we adopted a conservative approach assuming an event rate under  $H_0 = 0.5$  and an Odds ratio = 2.00, which also corresponds to our minimum effect size of interest. With a one-tailed  $\alpha = .05$ , and power = .80, the projected sample size needed was of approximately  $N = 217$ . More participants were collected than anticipated; due to an oversight, data collection continued after we had achieved the target sample size.

A total of 383 participants completed Experiment 1. This sample consisted of university students who participated in exchange for course credits and members of the general population recruited via advertisements on social media. During the data cleaning process, several exclusion criteria were applied to ensure data quality: (a) 25 cases were removed for not passing at least 4 out of 5 attention checks; (b) 2 cases were removed for completing the experiment in under 15 min (an impossible time to attentively complete the study), and (c) 23 cases were removed for taking more than 90 min to complete the experiment; (d) 8 cases were removed as participants self-reported having an English level lower than elementary; and (e) 3 cases were removed due to technical issues (e.g., video not playing). All exclusion criteria were defined a priori and pre-registered, with the main rationale of ensuring high-quality data for the main analysis. The final sample ( $N = 324$ ) had a mean age of  $M = 24.65$ , ranging from 18 to 71 years ( $SD = 8.13$ ) with 61% identifying as female. Most participants were recruited from the general population through social media (84%), followed by university students (16%). Students received course credits, and participants from social media were entered into a prize drawing for 10 Amazon vouchers with a value of €30.00 each.

### 2.1.3 | General judicial instruction video

The teaching aid variable was manipulated across three levels: general judicial instruction versus passive I-I-Eye aid versus active I-I-Eye aid. In the general judicial instruction condition, participants saw an animated video with generic information about a trial (5:42 min long), but with no details regarding eyewitness evidence. This animated educational video was created using the Vyond (2019) software. The content of the video is based on materials produced by Pawlenko et al., (2013) and was adapted to be applicable to a European legal context, thereby removing general legal information that was specific to certain jurisdictions. The general judicial instructions did not include any specific information about evaluating eyewitness evidence, instead it

contained legal definitions and general information about trials, emphasizing the importance of fairness and impartiality.

### 2.1.4 | I-I-Eye aid

The I-I-Eye aid is a three-step method for analyzing eyewitness evidence (Pawlenko et al., 2013; Safer et al., 2016). In this study, the I-I-Eye was presented in an animated educational video (8:32 min long), created using the Vyond (2019) software. The voice-over in the animated video uses the same content as the original I-I-Eye slides (Pawlenko et al., 2013). In the first segment of the video participants are instructed about important factors related to the eyewitness interview. The second segment covers important factors related to eyewitness identification procedures. The third and final segment includes information on factors related to the crime that can affect the eyewitness memory accuracy. The instructional video provides specific examples for all three segments. Participants are instructed that the accuracy of eyewitness testimonies should be questioned if the interview, identification procedure, or both, were unfairly conducted. Also, participants are informed that if the interview and identification procedures were conducted properly, the accuracy of an eyewitness testimony may still be acceptable even if the conditions of the crime were not ideal.

### 2.1.5 | I-I-Eye aid active component

An adapted version of a form proposed by Wise et al., 2009 was used as an active component complementing the I-I-Eye aid. In this form, participants were asked to evaluate the case at hand in relation to what they learned from the I-I-Eye aid. Participants were presented with a list of factors that might have biased the interview and identification procedures, as well as a list of crime factors that might have impaired eyewitness memory. For each factor, participants were prompted through an open-ended question to reflect upon whether the accuracy of the testimony was affected by the factors. Finally, participants were asked four questions to help them form a conclusion about the probable accuracy of the testimony (e.g., "Is there a high, medium, or low probability that the eyewitness identification was accurate?").

### 2.1.6 | Mock-trial videos

Videos of a simulated mock-trial about a robbery were adapted from Douglass and Jones (2013). As far as we are aware, this is the first time the I-I-Eye is being tested to evaluate the testimony of a victim in a robbery case, which is important for the generalizability of effects for this teaching aid. At the beginning of both videos, the judge provides information about the charges against the defendant and the burden of proof. The trial is then followed by direct, cross, and re-direct examination of the eyewitness, ending with the closing

arguments from both the defense and prosecution. The witness is the only person shown throughout the video. Two different versions of this video were created to manipulate for case strength. All of the changes in this manipulation were made in the direct examination by altering specific details of the crime (e.g., presence of a gun). In the weak case condition, the eyewitness states that: (a) the street lamp was broken so it was quite dark; (b) the perpetrator was holding a gun; (c) the perpetrator was wearing a hat; (d) the witness saw the perpetrator for about 10 s. In the strong case condition, the eyewitness states that: (a) it was very well lit; (b) no weapon is mentioned; (c) the perpetrator did not have a disguise; and (d) the witness saw the perpetrator for about 1 min. The opening statements, cross-examination, re-examination and closing statements are the same across both weak and strong conditions. Before the beginning of the trial, participants were instructed that the defendant: (a) was arrested because he matched the general description of the culprit given to police; (b) had no alibi for the time at which the robbery was said to have occurred; and (c) had a criminal record for armed robbery. These details were included in the trial phase because an initial pilot of the materials ( $N = 35$ ) showed little variance in verdict responses and an overly large proportion of not-guilty responses (85.7% not-guilty responses in the pilot). A second pilot after inclusion of incriminatory details showed a more balanced proportion of not-guilty responses (58.8%,  $N = 17$ ).

### 2.1.7 | Dependent measures and manipulation checks

The main dependent measure was the verdict (guilty/not guilty) provided by participants at the end of the mock-trial video. Participants were also asked to estimate how likely it is that the defendant was guilty on a scale from 1 - Extremely Unlikely to 7 - Extremely Likely. Manipulation checks were included to verify whether the weak and strong cases were indeed perceived as weak or strong (e.g., "Please rate the overall strength of the eyewitness testimony in helping the prosecution case"; 1 - Very Weak to 5 - Very Strong).

## 2.2 | Procedure

Experiment 1 was conducted in one online session via the survey platform Qualtrics. Participants were informed that they would be asked to evaluate a criminal trial, and that they would first watch a video containing important information required to evaluate the trial. At this stage, participants were randomly assigned to one of the three teaching aid conditions (general judicial aid vs. passive I-I-Eye aid vs. active I-I-Eye aid). Participants in the I-I-Eye active and I-I-Eye passive conditions watched the I-I-Eye teaching aid video, while the control condition watched the general judicial instructions. Next, participants were randomly assigned to watch one version of the mock-trial video (weak vs. strong). At the end of the trial, only participants in the I-I-Eye active aid condition received the active component form in which they were required to evaluate the eyewitness testimony in relation

to what they had learned from the I-I-Eye teaching aid. Following this, participants responded to the dependent measures, manipulation checks and demographic questions, receiving a debriefing at the end.

## 2.3 | Results

### 2.3.1 | Manipulation checks

Prior to the main analysis, ANOVAs were conducted to examine the manipulation checks, testing whether the differences between the weak and the strong cases were noticeable. The following dependent measures were compared across the two case strength conditions: the perceived strengths of the prosecutor's case, the defense's case and the eyewitness testimony. It was found that participants did not differ in their perceptions of the prosecution case in the weak ( $M = 3.05$ ,  $SD = 0.94$ ) and strong ( $M = 3.18$ ,  $SD = 0.95$ ) eyewitness evidence conditions,  $F(1,322) = 1.72$ ,  $p = .19$ ,  $\eta_p^2 = 0.05$ . However, participants perceived the defense case to be stronger in the weak eyewitness evidence condition ( $M = 2.60$ ,  $SD = 0.99$ ) compared to the strong eyewitness evidence condition ( $M = 2.39$ ,  $SD = 0.96$ ),  $F(1,322) = 3.74$ ,  $p = .05$ ,  $\eta_p^2 = 0.01$ . Similarly, participants perceived the eyewitness testimony to be stronger in the strong eyewitness evidence condition ( $M = 3.61$ ,  $SD = 0.87$ ) compared to the weak eyewitness evidence condition ( $M = 3.41$ ,  $SD = 0.93$ ),  $F(1,322) = 3.89$ ,  $p = .05$ ,  $\eta_p^2 = 0.01$ . Overall, it was deemed that the differences between the strong and weak eyewitness evidence conditions were somewhat noticeable, although such differences were small.

### 2.3.2 | Effects of teaching aids on verdicts

The main analysis was conducted on the verdict dependent variables using regression models to investigate whether the I-I-Eye teaching aids improved the evaluation of eyewitness evidence. Logistic regression models were fitted to test the factor variables teaching aid and case strength as predictors of the dichotomous verdict decision (guilty vs. not guilty). Three sets of orthogonal contrasts were calculated to examine how the teaching aids affected guilty verdicts for weak and strong eyewitness cases. The first contrast tested whether the teaching aids produced a difference in the guilty verdicts across all cases, examining whether the I-I-Eye aid increased skepticism (e.g., more not guilty verdicts for both strong and weak cases). The second and third contrasts were calculated to more specifically test how the teaching aid conditions affected sensitivity to the eyewitness evidence. The second contrast tested differences in guilty verdicts for each pair of teaching aid conditions (i.e., control vs. I-I-Eye passive; control vs. I-I-Eye active; and I-I-Eye passive vs. I-I-Eye active) when the eyewitness evidence was weak or strong (see Table 1). The third contrast tested differences in guilty verdicts for weak and strong eyewitness evidence for each teaching aid condition. To clarify our definition of sensitivity in the current paper, sensitivity is assessed in two different ways. The first approach to assess sensitivity is based on the differences in

**TABLE 1** Contrasts comparing guilty verdicts across teaching aid conditions and eyewitness evidence strength conditions (Experiment 1)

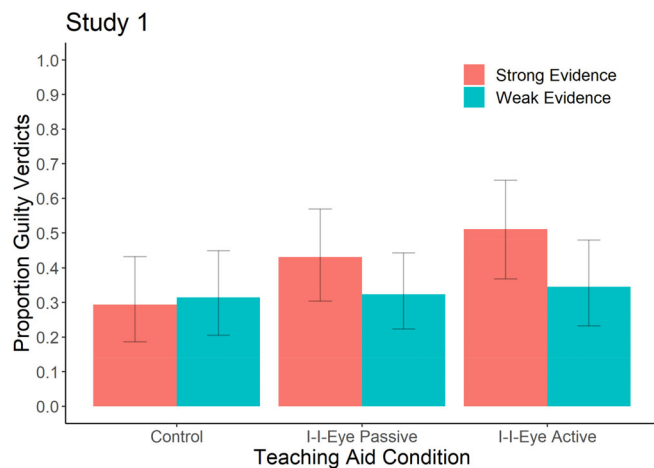
Contrast 1 (all cases)	%	OR	[95% CI]	<i>p</i>
Control	30.5%			
vs. I-I-Eye passive	37.0%	1.34	[0.76, 2.34]	.30
Control	30.5%			
vs. I-I-Eye active	42.0%	1.65	[0.93, 2.94]	.08
I-I-Eye passive	37.0%			
vs. I-I-Eye active	42.0%	1.23	[0.71, 2.13]	.44
Contrast 2 (split by case)				
Condition case = strong				
Control = 0	29.4%			
vs. I-I-Eye passive = 1	43.1%	1.82	[0.80, 4.13]	.15
Control = 0	29.4%			
vs. I-I-Eye active = 1	51.1%	2.51	[1.08, 5.81]	.03
I-I-Eye passive = 0	43.1%			
vs. I-I-Eye active = 1	51.1%	1.38	[0.61, 3.08]	.43
Condition case = weak				
Control = 0	31.5%			
vs. I-I-Eye passive = 1	32.4%	1.04	[0.48, 2.24]	.91
Control = 0	31.5%			
vs. I-I-Eye active = 1	34.5%	1.15	[0.52, 2.55]	.73
I-I-Eye passive = 0	32.4%			
vs. I-I-Eye active = 1	34.5%	1.10	[0.52, 2.34]	.79
Contrast 3 (split by teaching aid)				
Teaching aid condition = control				
Weak = 0	31.5%			
vs. strong = 1	29.4%	0.90	[0.39, 2.08]	.81
Teaching aid condition = I-I-Eye passive				
Weak = 0	32.4%			
vs. strong = 1	43.1%	1.58	[0.74, 3.36]	.23
Teaching aid condition = I-I-Eye active				
Weak = 0	34.5%			
vs. strong = 1	51.1%	1.98	[0.88, 4.44]	.09

Note: OR, Odds ratio; 0 indicates reference group.

verdicts within each condition, so that there should be significantly more guilty verdicts for cases with strong evidence than for cases with weak evidence. The second approach is based on a relative comparison of guilty/not-guilty verdicts in the I-I-Eye condition in relation to the control condition, so that the I-I-Eye conditions should yield significantly more guilty verdicts than control instructions for the strong case, and significantly fewer guilty verdicts than control instructions for the weak case.

The overall percentage of guilty verdicts was 30.5% for the control condition, 37.0% for the I-I-Eye passive condition and 42.0% for the I-I-Eye active condition. The first contrast showed that none of the teaching aid conditions produced skepticism (see Table 1). In fact, participants in the I-I-Eye active condition produced more guilty verdicts in relation to control condition, although this difference was not statistically significant (OR = 1.65, 95% CI = 0.93, 2.94,  $p = .08$ ). The

second contrast showed that participants in the I-I-Eye Active condition were 2.51 times more likely to produce a guilty verdict when the case was strong, compared to the Control condition (95% CI = 1.08, 5.81;  $p = .03$ ). This specific result is in line with the hypothesis that the I-I-Eye may increase sensitivity to eyewitness evidence. The third contrast indicated that participants were not able to discriminate their verdicts between the strong and weak case in any of the teaching aid conditions. For example, in the I-I-Eye Active condition, although the proportion of guilty verdicts was higher in the strong case (51.1%) than in the weak case (34.5%), this difference was not statistically significant ( $p = .09$ ). Finally, the general pattern of results indicated no significant differences in guilty verdicts between the I-I-Eye passive and I-I-Eye active conditions for weak (OR = 1.10, 95% CI = 0.52, 2.34,  $p = .79$ ) or strong (OR = 1.38, 95% CI = 0.61, 3.08,  $p = .43$ ) eyewitness evidence. Figure 1 presents an overview of these results



**FIGURE 1** Proportion of guilty verdicts per teaching aid condition and eyewitness evidence strength (Experiment 1)

with a plot of proportion of guilty verdicts per teaching aid and eyewitness evidence condition. In sum, this set of results indicate that although the I-I-Eye teaching aids did not create skepticism, there seems to be little evidence for improved sensitivity. More specifically, in Experiment 1 the only evidence in favor of the I-I-Eye improving sensitivity is the higher number of guilty verdicts in the I-I-Eye Active condition when the eyewitness evidence was strong.

## 2.4 | Discussion

In Experiment 1, we aimed to further test the effectiveness of the I-I-Eye and examine whether adding an active component to this teaching aid improves the evaluation of eyewitness testimony. We hypothesized that participants who received the I-I-Eye teaching aid (either the active or passive version) would be better able to distinguish weak and strong eyewitness evidence than a control group who received general judicial instructions. The results observed in Experiment 1 seem to provide some, although limited, support for this hypothesis. Crucially, it was found that none of the I-I-Eye teaching aids created skepticism, which would be flagged by a decrease in guilty verdicts across all types of eyewitness evidence. Instead, it was found that participants in the I-I-Eye teaching aid conditions tended to produce more guilty verdicts. The increase in guilty verdicts was specifically related to cases where eyewitness evidence was strong (i.e., it was well lit, no weapon present, no disguise and long exposure duration), showing that the I-I-Eye teaching aid was somewhat effective in increasing sensitivity to the eyewitness evidence for strong cases. This specific result replicates and provides support for previous findings that the I-I-Eye can increase sensitivity in evaluating eyewitness evidence (Pawlenko et al., 2013; Safer et al., 2016). However, we did not obtain any evidence that the I-I-Eye was effective in increasing non-guilty verdicts for weak cases, which raises some important concerns regarding the effectiveness of the I-I-Eye. That is because in order to safeguard against miscarriages of justice, it is particularly

important that teaching aids are able to increase non-guilty verdicts for weak cases. In the findings of Experiment 1, the number of guilty verdicts for weak cases did not seem to differ in any of the experimental conditions, averaging about 33% across all conditions. Furthermore, we have found no evidence that the active form in the I-I-Eye active condition provides a substantial improvement to the I-I-Eye aid, given that the general pattern of results was very similar for both conditions.

## 3 | EXPERIMENT 2

Experiment 1 provided some, although limited evidence that the I-I-Eye aid can increase sensitivity to eyewitness evidence strength. We also found no evidence that an active form in the I-I-Eye can improve the evaluation of eyewitness evidence. In Experiment 2 we aimed to further test our main hypotheses with a new independent sample and a slightly modified method. More specifically, in Experiment 2 we investigate whether the I-I-Eye Passive and I-I-Eye Active aids can improve the sensitivity of laypeople to eyewitness evidence when the differences between weak and strong cases are more salient. While in Experiment 1 the main manipulation of weak vs. strong cases consisted of changing four details in the eyewitness evidence, in Experiment 2 the weak versus strong cases were manipulated by changing 10 details in the eyewitness evidence. These 10 details also had a higher focus on system variables rather than estimator variables, in order for the results of Experiment 2 to be complementary to the findings of Experiment 1 and to make Experiment 2 more comparable to Safer et al., (2016). It was argued that this new manipulation could help clarify whether the I-I-Eye teaching aid and its active form component have a beneficial effect on the evaluation of eyewitness evidence, given that if a beneficial effect exists then it should be more easily observed when there are more noticeable discrepancies between weak and strong eyewitness evidence. Another reason for following up on this new manipulation can be found in the data of Experiment 1, given that some results regarding the manipulation checks showed that the weak and strong cases were only somewhat noticeable. All other methodological aspects of Experiment 1 were kept constant in Experiment 2 given that a meta-analysis combining the effects from both experiments was also planned to obtain estimates that are more precise and that take into account a wider range of eyewitness evidence cases.

### 3.1 | Methods

#### 3.1.1 | Design

Experiment 2 had the same design as Experiment 1: a 3 (teaching aid: general judicial instruction vs. passive I-I-Eye aid vs. active I-I-Eye aid)  $\times$  2 (eyewitness evidence: strong vs. weak case) between-subjects factorial design. Participants were randomly allocated to one out of the six experimental conditions.

### 3.1.2 | Participants

A total of 346 participants completed this experiment. This sample consisted of workers from Prolific who participated in exchange for £4.85. During the data cleaning process, several exclusion criteria were applied to ensure data quality: (a) 20 cases were removed for not passing at least 4 out of 5 attention check; (b) 1 case was removed as participants self-reported having an English level lower than elementary; and (c) 3 cases were removed due to technical issues (e.g., video not playing). The final sample ( $N = 322$ ) had a mean age of  $M = 30.00$ , ranging from 18 to 70 years ( $SD = 11.70$ ) and included 36% female participants.

### 3.1.3 | Teaching aids

All details regarding the General Judicial Aid and I-I-Eye teaching aids were the exact same as in Experiment 1, including the details regarding the active form in the I-I-Eye active condition.

### 3.1.4 | Mock trial video

The main difference between Experiment 1 and Experiment 2 was in the mock trial video, given that in Experiment 2 the weak and strong

eyewitness cases differed in 10 points, while in Experiment 1 the weak and strong eyewitness cases differed in 4 points. Experiment 2 used the same general structure of a simulated mock-trial about a robbery adapted from Douglass and Jones (2013). Two different versions of this video were created to manipulate for case strength. All of the changes in this manipulation were made in the direct examination by altering specific details of the crime (see Table 2; details were adapted from Safer et al., 2016). The opening statements, cross-examination, re-examination and closing statements are the same across both weak and strong conditions. Before the beginning of the trial, participants were instructed that the defendant: (a) was arrested because he matched the general description of the culprit given to police; (b) had no alibi for the time at which the robbery was said to have occurred; and (c) had a criminal record for armed robbery.

### 3.1.5 | Dependent measures

Similarly to Experiment 1, the main dependent measure was the verdict (guilty/not guilty) provided by participants at the end of the mock-trial video. Participants were also asked to estimate how likely it is that the defendant was guilty on a scale from 1 - Extremely Unlikely to 7 - Extremely Likely. Manipulation checks were also included to verify whether the weak and strong case were indeed perceived as weak or strong (e.g., "Please rate the overall strength of the

**TABLE 2** List of details manipulated for strong and weak eyewitness cases

Manipulated details	Strong case	Weak case
Whether the interviewing officer asked the eyewitness about media exposure. [I]	Officer asked if witness had read or seen news reports about the crime.	Officer did not ask if witness had read or seen news reports about the crime.
Instruction to the eyewitness to avoid discussing the crime and avoid media stories [I]	Witness was asked not to talk about what she saw with anyone other than the police.	Witness watched some news reports about the crime and discussed the event with family and friends.
Leading interview questions about the color of the perpetrator's hair and the quality of the eyewitness's view of the perpetrator [I]	No leading questions during eyewitness interview.	Leading questions during eyewitness interview.
Standardized lineup instructions. [I]	Eyewitness: was told she would see photographs of several men and that she would need to say "yes" if that was the man she saw committing the crime, or "no" if it wasn't.	Eyewitness was told she would be seeing a set of photographs and then choose the man who looked most familiar.
Number of lineup members. [L]	Eight photos in the lineup.	Five photos in the lineup.
Blind lineup administration. [L]	The officer showing the lineup did not know who the suspect was.	The officer showing the lineup was aware of who the suspect was.
Description-matched lineup. [L]	All fillers matched the description of the perpetrator.	Fillers did not match the description of the perpetrator
Cautionary instruction that the perpetrator may or may not be in the lineup. [L]	Witness was told that the culprit may or may not be in the lineup.	Witness was not instructed that the culprit may or may not be in the lineup.
Statement of confidence taken immediately after the eyewitness's identification. [L]	Confidence was collected immediately after the identification.	Confidence was not collected immediately after the identification.
Confirming feedback immediately after the identification. [L]	Eyewitness did not receive confirmatory feedback after the identification.	Eyewitness was told by the officer that she had identified their suspect and that she did a good job.

Note: [I] Indicates details related to the interview; [L] Indicates details related to the lineup identification task.



eyewitness testimony in helping the prosecution case"; 1 - Very Weak to 5 - Very Strong).

## 3.2 | Procedure

The procedure in Experiment 2 was the same as in Experiment 1. Experiment 2 was conducted in one online session via the survey platform Qualtrics. Participants were informed that they would be asked to evaluate a criminal trial, and that they would first watch a video containing important information required to evaluate the trial. At this stage participants were randomly assigned to one of the three teaching aid conditions (general judicial aid vs. passive I-I-Eye aid vs. active I-I-Eye aid). Participants in the I-I-Eye active and I-I-Eye passive conditions watched the I-I-Eye teaching aid video, while the control condition watched the General Judicial Instructions. Next, participants were randomly assigned to watch one version of the mock-trial video (weak vs. strong). At the end of the trial, only participants in the I-I-Eye Active aid condition received the active form in which they were required to evaluate the eyewitness testimony in relation to what they have learned from the I-I-Eye teaching aid. Next, participants responded to the dependent measures, manipulation checks and demographic questions, receiving a debriefing in the end.

## 3.3 | Results

### 3.3.1 | Manipulation checks

Prior to the main analysis, ANOVAs were conducted to examine the manipulation checks, testing whether the differences between the weak and the strong cases were noticeable. The following dependent measures were compared across the two case strength conditions: the perceived strength of the prosecutor's case, the defense's case and the eyewitness testimony. Participants found the prosecution's case to be stronger in the strong eyewitness evidence condition ( $M = 3.55$ ,  $SD = 0.89$ ) compared to the weak eyewitness evidence condition ( $M = 3.24$ ,  $SD = 0.96$ ),  $F(1,320) = 8.66$ ,  $p = .003$ ,  $\eta_p^2 = 0.03$ . Similarly, participants found the defense case to be stronger in the weak eyewitness evidence condition ( $M = 3.28$ ,  $SD = 0.92$ ) compared to the strong eyewitness evidence condition ( $M = 2.68$ ,  $SD = 1.03$ ),  $F(1,320) = 31.47$ ,  $p < .001$ ,  $\eta_p^2 = 0.09$ . Furthermore, participants found the eyewitness testimony to be stronger in the strong eyewitness evidence condition ( $M = 3.86$ ,  $SD = 0.91$ ) compared to the weak eyewitness evidence condition ( $M = 3.39$ ,  $SD = 1.08$ ),  $F(1,320) = 17.92$ ,  $p < .001$ ,  $\eta_p^2 = 0.05$ . Overall, it was deemed that the differences between the strong and weak eyewitness evidence conditions were considerably noticeable.

### 3.3.2 | Effects of teaching aids on verdicts

The main analysis was conducted on the verdict dependent variables using regression models to investigate whether the I-I-Eye teaching

aids improved the evaluation of eyewitness evidence. Logistic regression models were fitted to test the factor variables teaching aid and case strength as predictors of the dichotomous verdict decision (guilty versus not guilty). Three sets of orthogonal contrasts were calculated to examine how the teaching aids affect guilty verdicts for weak and strong eyewitness cases. The described contrasts were defined a priori and pre-registered to test for our specific hypotheses. The first contrast tested whether the teaching aids produced a difference in the guilty verdicts across all cases, examining whether the I-I-Eye aid increased skepticism (e.g., more not guilty verdicts for both strong and weak cases). The second and third contrast were calculated to more specifically test how the teaching aid conditions affected sensitivity to the eyewitness evidence. The second contrast tested differences in guilty verdicts for each pair of teaching aid conditions (i.e., control vs. I-I-Eye passive; control vs. I-I-Eye active; and I-I-Eye passive vs. I-I-Eye active) when the eyewitness evidence was weak or strong (see Table 3). The third contrast tested differences in guilty verdicts for weak and strong eyewitness evidence for each teaching aid condition.

In Experiment 2, the overall percentage of guilty verdicts was 43.4% for the Control condition, 45.7% for the I-I-Eye Passive condition and 47.1% for the I-I-Eye Active condition. The first contrast showed that none of the teaching aid conditions produced skepticism (see Table 3). The second contrast showed that participants in the I-I-Eye Passive condition were 2.15 times more likely to produce a guilty verdict when the case was strong, compared to the Control condition (95% CI = 0.99, 4.66;  $p = .03$ ). The third contrast indicated that participants were able to discriminate between the strong and weak case in the teaching aid conditions, with a more pronounced effect in the I-I-Eye Passive condition. Participants in the I-I-Eye passive condition were 3.64 times more likely to give a guilty verdict when the case was strong, compared to when it was weak (95% CI = 1.62, 8.18;  $p = .001$ ). Participants in the I-I-Eye active condition were 2.16 times more likely to give a guilty verdict when the case was strong, compared to when it was weak (95% CI = 0.98, 4.73;  $p = .05$ ). Participants in the Control condition produced a very similar number of guilty verdicts in the weak (43.6%) and strong cases (43.1%), OR = 0.98,  $p = .95$ .

No significant differences in guilty verdicts were found between the I-I-Eye Passive and I-I-Eye Active conditions for weak (OR = 1.38, 95% CI = 0.62, 3.04,  $p = .42$ ) or strong (OR = 0.81, 95% CI = 0.36, 1.83,  $p = .62$ ) eyewitness evidence. Figure 2 presents an overview of these results with a plot of proportion of guilty verdicts per teaching aid and eyewitness evidence condition. In sum, this set of results indicate that the I-I-Eye teaching aids did improve eyewitness sensitivity to eyewitness evidence, although no differences were observed between the I-I-Eye passive and I-I-Eye active condition.

### 3.3.3 | Meta-analysis

Given that Experiment 1 and Experiment 2 presented, in general, a similar pattern of results, a meta-analysis of these two experiments was conducted to obtain estimates that are more precise and that account for a wider variety of eyewitness cases (Goh et al., 2016).

**TABLE 3** Contrasts comparing guilty verdicts across teaching aid conditions and eyewitness evidence strength conditions (Experiment 2)

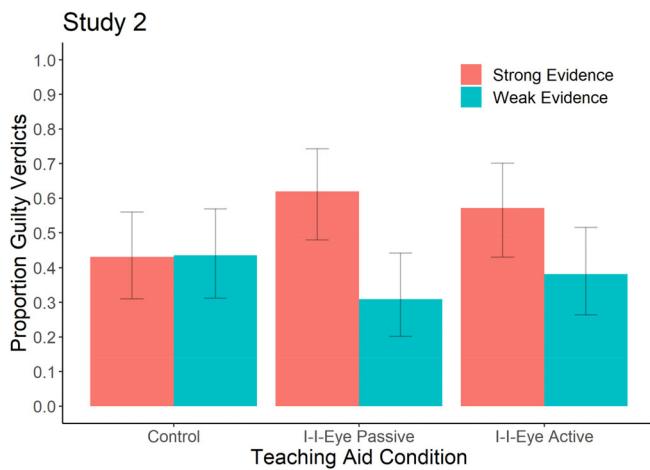
Contrast 1 (all cases)	%	OR	[95% CI]	<i>p</i>
Control	43.4%			
vs. I-I-Eye passive	45.7%	1.10	[0.64, 1.88]	.72
Control	43.4%			
vs. I-I-Eye active	47.1%	1.16	[0.68, 1.99]	.58
I-I-Eye passive	45.7%			
vs. I-I-Eye active	47.1%	1.06	[0.61, 1.82]	.84
Contrast 2 (split by case)				
Condition case = strong				
Control = 0	43.1%			
vs. I-I-Eye passive = 1	62.0%	2.15	[0.99, 4.66]	.05
Control = 0	43.1%			
vs. I-I-Eye active = 1	57.1%	1.76	[0.81, 3.79]	.14
I-I-Eye passive = 0	62.0%			
vs. I-I-Eye active = 1	57.1%	0.81	[0.36, 1.83]	.62
Condition case = weak				
Control = 0	43.6%			
vs. I-I-Eye passive = 1	30.9%	0.58	[0.26, 1.26]	.17
Control = 0	43.6%			
vs. I-I-Eye active = 1	38.2%	0.79	[0.37, 1.71]	.56
I-I-Eye passive = 0	30.9%			
vs. I-I-Eye active = 1	38.2%	1.38	[0.62, 3.04]	.42
Contrast 3 (split by teaching aid)				
Teaching aid condition = control				
Weak = 0	43.6%			
vs. strong = 1	43.1%	0.98	[0.46, 2.06]	.95
Teaching aid condition = I-I-Eye passive				
Weak = 0	30.9%			
vs. strong = 1	62.0%	3.64	[1.62, 8.18]	.001
Teaching aid condition = I-I-Eye active				
Weak = 0	38.2%			
vs. strong = 1	57.1%	2.16	[0.98, 4.73]	.05

Note: OR, Odds ratio; 0 indicates reference group.

This analysis was conducted using the meta package in R (Harrer et al., 2019; Schwarzer, 2007). The meta-analysis consisted of pooling the effect sizes (i.e., odds ratio) from the main logistic regression models and contrasts as outlined in both experiments. A random-effects model was used in the meta-analysis given that it can be expected that the population of studies in this topic area will have considerable between-study heterogeneity in the true effects. The variance of the distribution of true effect sizes ( $\tau^2$ ) was estimated using the Paule-Mandel estimator, as it is considered a more appropriate estimator for binary effect size data (Harrer et al., 2019).

In the meta-analysis ( $N = 646$ ), the overall percentage of guilty verdicts was 37.2% for the Control condition, 41.1% for the I-I-Eye Passive condition and 44.6% for the I-I-Eye Active condition. The meta-analysis of the first contrast showed that none of the teaching aid conditions produced skepticism (see Table 4). The meta-analysis of the second contrast showed that participants in the I-I-Eye Passive condition were 1.99

times more likely to produce a guilty verdict when the case was strong, compared to the Control condition (95% CI = 1.13, 3.49;  $p = .02$ ). Moreover, the second contrast showed that participants in the I-I-Eye Active condition were 2.07 times more likely to produce a guilty verdict when the case was strong, compared to the Control condition (95% CI = 1.17, 3.64,  $p = .01$ ). The meta-analysis of the third contrast indicated that participants were able to discriminate between the strong and weak case in the teaching aid conditions, with a more pronounced effect in the I-I-Eye Passive condition. Participants in the I-I-Eye Passive condition were 2.37 times more likely to give a guilty verdict when the case was strong, compared to when it was weak (95% CI = 1.04, 5.36;  $p = .04$ ). Participants in the I-I-Eye Active condition were 2.07 times more likely to give a guilty verdict when the case was strong, compared to when it was weak (95% CI = 1.18, 3.63;  $p = .01$ ). Participants in the control condition produced a very similar number of guilty verdicts in the weak (36.7%) and strong cases (37.6%), OR = 0.95,  $p = .84$ .



**FIGURE 2** Proportion of guilty verdicts per teaching aid condition and eyewitness evidence strength (Experiment 2)

Contrast 1 (all cases)	%	OR	[95% CI]	<i>p</i>
Control	37.2%			
vs. I-I-Eye passive	41.1%	1.20	[0.82, 1.77]	.33
Control	37.2%			
vs. I-I-Eye active	44.6%	1.36	[0.92, 2.02]	.11
I-I-Eye passive	41.1%			
vs. I-I-Eye active	44.6%	1.14	[0.78, 1.67]	.49
Contrast 2 (split by case)				
Condition case = strong				
Control = 0	36.7%			
vs. I-I-Eye passive = 1	52.5%	1.99	[1.13, 3.49]	.02
Control = 0	36.7%			
vs. I-I-Eye active = 1	54.3%	2.07	[1.17, 3.64]	.01
I-I-Eye passive = 0	52.5%			
vs. I-I-Eye active = 1	54.3%	1.06	[0.60, 1.87]	.83
Condition case = weak				
Control = 0	37.6%			
vs. I-I-Eye passive = 1	31.7%	0.78	[0.44, 1.39]	.39
Control = 0	37.6%			
vs. I-I-Eye active = 1	36.4%	0.95	[0.55, 1.64]	.85
I-I-Eye passive = 0	31.7%			
vs. I-I-Eye active = 1	36.4%	1.23	[0.71, 2.12]	.46
Contrast 3 (split by teaching aid)				
Teaching aid condition = control				
Weak = 0	36.7%			
vs. strong = 1	37.6%	0.95	[0.55, 1.65]	.84
Teaching aid condition = I-I-Eye passive				
Weak = 0	31.7%			
vs. strong = 1	52.5%	2.37	[1.04, 5.36]	.04
Teaching aid condition = I-I-Eye active				
Weak = 0	36.4%			
vs. strong = 1	54.3%	2.07	[1.18, 3.63]	.01

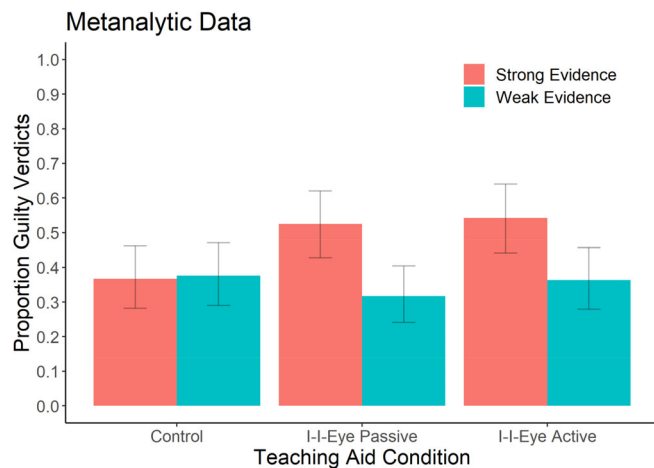
Note: OR, Odds ratio; 0 indicates reference group.

No significant differences in guilty verdicts were found between the I-I-Eye Passive and I-I-Eye Active conditions for weak (OR = 1.23, 95% CI = 0.71, 2.12,  $p = .46$ ) or strong (OR = 1.06, 95% CI = 0.60, 1.87,  $p = .83$ ) eyewitness evidence. Figure 3 presents an overview of these results with a plot of proportion of guilty verdicts per teaching aid and eyewitness evidence condition. In sum, results from the meta-analysis indicate that the I-I-Eye teaching aids did improve eyewitness sensitivity to eyewitness evidence, although no differences were observed between the I-I-Eye passive and I-I-Eye active condition.

### 3.4 | Exploratory analysis

Some exploratory analyses were also conducted by examining the effect of our main manipulations (i.e., Control, I-I-Eye passive and I-I-Eye active) on other dependent measures. The dependent measures

**TABLE 4** Contrasts comparing guilty verdicts across teaching aid conditions and eyewitness evidence strength conditions (meta-analysis)



**FIGURE 3** Proportion of guilty verdicts per teaching aid condition and eyewitness evidence strength (metanalytic data)

explored were: (a) the continuous verdict (ranging from 1 - Extremely Unlikely to 7 - Extremely Likely; higher scores indicative of more likelihood of guilt); (b) the perceived strength of the prosecution case; (c) the perceived strength of the defense case; and (d) the perceived strength of the eyewitness testimony (all ranging from 1 - Very Weak to 5 - Very Strong). The main rationale for this exploratory analysis was to further investigate the effect of receiving the I-I-Eye on jurors' sensitivity to eyewitness evidence.

A series of four regression models were fitted by using the I-I-Eye conditions and Case Strength conditions as predictor variables and the dependent variables as predicted variables. An interaction term was also included between I-I-Eye conditions and Case Strength conditions. The general pattern of results seemed to confirm the main findings observed in Experiment 1 and Experiment 2 (see Supplemental Materials S1 for full report of the results). For example, it was found that in Experiment 1, there were no significant differences in the continuous verdict measure between weak and strong cases in the control condition ( $M_{\text{weak}} = 5.00$ ,  $M_{\text{strong}} = 5.22$ ,  $\beta = 0.21$ ,  $p = .38$ ), I-I-Eye passive condition ( $M_{\text{weak}} = 5.24$ ,  $M_{\text{strong}} = 5.47$ ,  $\beta = 0.23$ ,  $p = .31$ ) and I-I-Eye active condition ( $M_{\text{weak}} = 4.82$ ,  $M_{\text{strong}} = 5.31$ ,  $\beta = 0.49$ ,  $p = .05$ ). However, in experiment two (i.e., where eyewitness issues were more noticeable), the results showed no differences in the continuous verdict measure between weak and strong cases in the Control condition ( $M_{\text{weak}} = 5.04$ ,  $M_{\text{strong}} = 4.84$ ,  $\beta = -0.19$ ,  $p = .47$ ), but there were significant differences in continuous verdict between weak and strong cases in the I-I-Eye Passive condition ( $M_{\text{weak}} = 4.49$ ,  $M_{\text{strong}} = 5.64$ ,  $\beta = 1.14$ ,  $p < .001$ ) and I-I-Eye Active condition ( $M_{\text{weak}} = 4.55$ ,  $M_{\text{strong}} = 5.31$ ,  $\beta = 0.49$ ,  $p < .001$ ). These results show that participants receiving the I-I-Eye found the defendant more likely to be guilty when the eyewitness evidence was strong compared to when it was weak. Participants in the control condition gave similar guilty verdict likelihoods for weak and strong eyewitness evidence. This same pattern of results can be observed for the other dependent measures (i.e., the perceived strength of the prosecution case; the perceived strength of the defense case; and the perceived strength of the

eyewitness testimony). Full statistical parameters are presented in the Supplemental Materials S1.

## 4 | GENERAL DISCUSSION

The main goal of the present study was to further investigate the effectiveness of the I-I-Eye and examine whether adding an active component to this teaching aid would improve the evaluation of eyewitness testimony. We hypothesized that participants who received the I-I-Eye teaching aid, either in the active or passive format, would be able to better distinguish weak from strong eyewitness evidence compared to a control condition. Participants in the control condition received general judicial instructions without the mentioning of eyewitness evidence issues. Given that Experiment 1 and Experiment 2 had a similar pattern of results in general, we will tailor our main inferences and discussion to the results found in the meta-analysis of both experiments, highlighting important differences between both experiments.

In line with previous findings, Experiment 1 and 2 concurrently demonstrated that the I-I-Eye improved jurors' sensitivity to eyewitness evidence in strong cases. In the meta-analysis, compared with the control participants the I-I-Eye participants in the passive and active condition were approximately two times more likely to choose guilty verdicts for the strong case rather than for the weak case. In other words, participants in the I-I-Eye conditions were better able to discriminate strong eyewitness evidence from weak eyewitness evidence, compared to participants who only received general judicial instructions. However, an important shortcoming in our findings is that the I-I-Eye did not seem to be as effective in situations where the eyewitness evidence is weak. In other words, it was found that jurors receiving the I-I-Eye may find the defendant more likely to be guilty if the eyewitness evidence is strong, but if the eyewitness evidence is weak the I-I-Eye did not seem to effectively decrease perceptions of guilty. This is an important limitation given that for the I-I-Eye to be fully effective it would be expected for it to also decrease perceptions of guilty when the eyewitness evidence is weak, in order to avoid potential wrongful convictions. Overall, our findings indicate some mixed evidence regarding the effectiveness of the I-I-Eye. We now discuss how our findings relate to other findings in the literature and potential factors in our experiments that may have contributed to this pattern of results.

Albeit still limited in scope, the literature around the I-I-Eye has been showing some promising findings, typically demonstrating that this teaching aid is an effective tool for improving the evaluation of eyewitness evidence (Pawlenko et al., 2013; Safer et al., 2016). An overview of the literature findings reveals that, when compared to control conditions, participants receiving the I-I-Eye give roughly 6%–25% more guilty verdicts in cases with strong eyewitness evidence, and roughly 14%–20% less guilty verdicts in cases with weak eyewitness evidence (i.e., Pawlenko et al., 2013; Safer et al., 2016). More recently, it was also found that the I-I-Eye can increase the effectiveness of expert testimony, as shown in a higher sensitivity

within an expert I-I-Eye condition compared to a standard expert condition (Wise & Kehn, 2020). The current study only partially confirms this general pattern of findings. The findings that support the general trend of effectiveness of the I-I-Eye are based on the increase of guilty verdicts in strong eyewitness cases, showing that participants were more trustworthy of the eyewitness evidence if there were no reasons to believe it was harmed by adverse eyewitness factors. However, it was expected that the I-I-Eye would also increase non-guilty verdicts for weak eyewitness cases, which was not observed in the current experiments.

One important distinction across different investigations on the I-I-Eye has to do with the amount and types of eyewitness factors being manipulated between weak and strong cases, as well as additional incriminatory evidence in the case. Some studies have manipulated 11 different factors (Pawlenko et al., 2013; Safer et al., 2016), while others have manipulated as many as 21 different factors (Wise & Kehn, 2020), also showing great variability in the type of factors being manipulated (e.g., estimator variables or system variables). In the current study, Experiment 1 included the manipulation of four eyewitness factors between weak and strong cases (e.g., weapon presence, exposure duration, light conditions, disguise), while Experiment 2 included 10 differences in the strong and weak cases. The pattern of results in both experiments was generally similar, with a seemingly more pronounced effectiveness of the I-I-Eye in Experiment 2. Similar patterns of results have also been observed in previous studies (Pawlenko et al., 2013; Safer et al., 2016), suggesting that the I-I-Eye had a larger effect in situations involving a higher number of eyewitness factors to be considered. One possible explanation as to why the current experiments do not produce the exact same pattern of results from previous findings may be due to the specific eyewitness factors being manipulated. The main argument against this explanation though is that Experiment 2 has tested a very similar list of manipulations as in previous studies (i.e., Pawlenko et al., 2013; Safer et al., 2016). So it may be more reasonable to assume that differences in findings can be attributed to other aspects of methodology, such as the accompanying incriminatory evidence presented on trial (e.g., alibi or fingerprint). It is also worth noting that Experiment 1 focused specifically on estimator variables (e.g., lighting, weapon focus), while Experiment 2 focused on system variables (e.g., lineup and interviewing procedures). Therefore, some of the differences between findings of Experiment 1 and Experiment 2 may also be attributed to the type of witnessing factors being examined. It is therefore crucial that further research is produced on scenarios where the I-I-Eye may be more or less effective, depending for example on the amount and type of eyewitness factors being discussed during trial. The practical implications of teaching aids may only be possible after careful examination of such varying scenarios, and future meta-analytical effects may be better able to clarify the general effectiveness of the I-I-Eye.

It is still noteworthy to point out that overall findings of the I-I-Eye seem to be in sharp contrast to what is observed for other types of teaching aids and judicial instructions. Many studies have now found that other instructions that are meant to sensitize jurors to

eyewitness issues such as the Henderson Instructions (New Jersey v. Henderson, 2011) or the Biggers factors (Neil v. Biggers, 1972) generally creates an overall skepticism of eyewitness evidence (Dillon et al., 2017; Jones et al., 2017; Papailiou et al., 2015). That is, when exposed to these instructions, jurors tend to disregard all types of eyewitness evidence, even when there is reason to believe that the eyewitness testimony is strong (e.g., when witnessing conditions, identification procedures and interviewing procedures are adequate). Therefore, the I-I-Eye seems to be showing an advantage in terms of its applied value and practical use, although the reason for its superiority is not entirely clear. It could be that the comprehensive list of factors covered in the I-I-Eye contributes to its effectiveness, given that other types of instructions tend to contain very little information on eyewitness issues (e.g., Biggers factors; Bradfield & Wells, 2000). A counter-argument to this hypothesis is that the seemingly ineffective Henderson instructions (New Jersey v. Henderson, 2011) also covers a wide range of eyewitness factors, but still seems to generate skepticism (Dillon et al., 2017; Jones et al., 2017). It could then be proposed that the I-I-Eye's effectiveness arises from its general structure, in terms of distinctively dividing its content into interviewing, identification and eyewitness factors. In sum, the exact reason why the I-I-Eye seems to be more effective than its counterparts is still not well understood, so further research in this area would be encouraged to clarify the I-I-Eye's advantages and to further improve its effectiveness.

#### 4.1 | The effectiveness of different I-I-Eye formats (active vs. passive)

One of the main goals of the current study was to test whether adding an active component to the I-I-Eye would further improve jurors' sensitivity to eyewitness evidence. The rationale for this investigation was based on theoretical frameworks that predict better retention when individuals analyze, synthesize, and evaluate educational materials rather than just passively receiving them (Chi & Wylie, 2014; Fredricks et al., 2004). It was therefore expected that jurors would be better able to integrate the I-I-Eye knowledge into their decision making by completing an active form component where jurors would be prompted to actively engage with the material by integrating and applying the I-I-Eye knowledge they had just learnt to the case at hand. However, neither results from Experiment 1 nor Experiment 2 showed significant differences when comparing the performance of the active and passive I-I-Eye. The evidence suggests that participants were better able to discriminate between weak and strong evidence when presented with either the active or passive I-I-Eye, compared to the control condition. However, there was no evidence that the active form provided any additional benefit to the sensitization of jurors to eyewitness evidence. More specifically, in the meta-analysis, it was found that the percentage of guilty verdicts was rather similar for the I-I-Eye passive and I-I-Eye active conditions in strong cases (52.5% and 54.3%, respectively) and weak cases (31.7% and 36.4%, respectively).

A few different explanations may be proposed for this unexpected result. Foremost, it may be the case that the I-I-Eye in its

original format already achieves high standards of sensitization for eyewitness evidence, in which case any additional components could confer little to no improvement in sensitization. It could also be argued that the particular format chosen for the active form (listing factors that might have impaired eyewitness memory or biased the interview and identification procedures) may have been ineffective or may have caused undue burden on jurors, who needed to verify all potential witness issues in the case at hand (e.g., heightened cognitive load). This broad format for the active form was chosen in order to test a component that could potentially be applied to any case involving eyewitness evidence, despite the presence or absence of certain eyewitness factors.

Another important explanation for the absence of an effect for the active versus passive I-I-Eye versions is that perhaps Experiment 1 and Experiment 2 may have been underpowered to detect such an effect. That is because target sample sizes were estimated by comparing the I-I-Eye to a zero-sensitive control group. However, if the I-I-Eye Passive version already achieves some sensitivity, then a greater number of cases would be needed to demonstrate that the I-I-Eye Active generates more sensitivity than the I-I-Eye Passive. Nevertheless, the current findings do not point to any evidence that the I-I-Eye Active is indeed worth implementing. Although there could still be circumstances where active components may improve the I-I-Eye's effectiveness (e.g., complex cases with various witnessing factors or multiple testimonies), it seems evident that the findings in the current study discourages further pursuit of this specific avenue of improvement for the I-I-Eye. It seems reasonable, however, that future studies could explore other potential improvements for the I-I-Eye, either in its content or presentation format.

## 4.2 | I-I-Eye in video format

One particularly important contribution of the current study is that it is the first to adapt and test the I-I-Eye in a video format, as opposed to the original format of Powerpoint slides (Pawlenko et al., 2013; Safer et al., 2016). One of the main rationales for adapting the I-I-Eye to a video format was to make this teaching aid more engaging, especially when used in an online setting. The I-I-Eye in video format presents the same content as the original I-I-Eye, although it makes use of animations and visual effects to present the content in a more engaging and appealing manner. Therefore, the current study presents some important initial findings regarding the effectiveness of the I-I-Eye in such video format, given that evidence in favor of sensitization for eyewitness factors was observed in Experiments 1 and 2.

An important implication of the findings in favor of the I-I-Eye video effectiveness is that courts might be more likely to implement a short instructional video than to undertake the role of classroom instructor and PowerPoint presenter (Safer et al., 2016). Moreover, short videos have been shown to be effective teaching tools in a variety of populations (Brame, 2016; Carmichael et al., 2018). Studies in this area have shown that short videos can help clarify complex concepts, promote further engagement, and enhance learning experiences

(Cherrett et al., 2009). If future studies are able to provide further support for the I-I-Eye video effectiveness, implementing this method into courtrooms and using it in regular practice could be considered, as it is an easy tool to use and gives other legal professionals an easier, common method for evaluating testimonies (Safer et al., 2016). A short instructional video used in a courtroom setting could be more cost and time effective compared to content being presented by an instructor, with the added benefit of ensuring that triers of fact are receiving the same information in the same manner. One limitation of the current study is that it did not aim to directly compare the video and Powerpoint slides format of the I-I-Eye, given the focus on testing the effectiveness of an active component. It is also important to highlight once again that the lack of sensitivity of the I-I-Eye for weak cases observed in the current study should be a reason of skepticism for the effectiveness of the I-I-Eye in video format. Thus, the current study provides limited evidence that the I-I-Eye is effective in its video format, making it crucial that further investigations are pursued to specifically test the effectiveness of the video format in comparison with the original I-I-Eye.

## 5 | LIMITATIONS

The current study presents a number of limitations. First, although some important methodological elements were explored further in Experiment 2, such as the type of eyewitness factors included, some other methodological elements were kept consistent in both experiments and therefore are underexplored. For example, so that the results from two experiments would be comparable, the same witness was shown in both Experiment 1 and 2, therefore making the results less generalizable to situations involving different witnesses. The perceived credibility of eyewitnesses has been found to be dependent on witness's characteristics such as age and likeability (Mueller-Johnson et al., 2007), raising doubts as to whether the I-I-Eye can be equally effective for situations involving different witnesses. Another important limitation is that the current study was conducted in an online setting using a survey format, conditions that are radically different from those present in a courtroom trial. Various important elements of court proceedings such as expert testimonies and jury deliberation were omitted from the current study, so it is important for the current findings to be interpreted in light of these limitations. Nevertheless, the current study provides some important evidence as to the effects of the I-I-Eye in jury decision-making across a variety of conditions. We recommend that future studies expand on these findings by testing situations with different witnesses and testing the I-I-Eye effectiveness under more comprehensive trial scenarios.

Another important limitation in our findings is that the I-I-Eye did not seem to be as effective in situations where the eyewitness evidence is weak. In other words, it was found that jurors receiving the I-I-Eye may find the defendant more likely to be guilty if the eyewitness evidence is strong, but if the eyewitness evidence is weak the I-I-Eye did not seem to effectively decrease perceptions of guilty. This

is an important limitation given that for the I-I-Eye to be fully effective it would be expected for it to also decrease perceptions of guilty when the eyewitness evidence is weak, in order to avoid potential wrongful convictions. This specific pattern of results need to be further analyzed and tested in future studies to confirm whether the I-I-Eye fails to achieve this effectiveness as a safeguard to weak eyewitness evidence. One potential explanation for these findings may have to do with the amount of incriminatory evidence against the defendant. To achieve more variability in verdict rates, participants in our studies were informed that the defendant was arrested because he matched the general description of the culprit, had no alibi, and had a criminal record for armed robbery. The inclusion of this incriminating, non-eyewitness evidence deliberately made the weak evidence condition less weak, so that there was a 31%–44% conviction rate for weak evidence in all conditions in the two experiments. Some evidence shows that jurors may give a guilty verdict with as low as 60% certainty of guilt, (Magnussen et al., 2014). Therefore, it may be the case that unless jurors are dealing with an extremely weak case, there is always going to be some jurors reaching a guilty verdict regardless of quality of the eyewitness evidence or educational aids. In practice, this means that the I-I-Eye must be further tested in situations involving other types of incriminatory evidence, to further conclude whether this teaching aid can still be beneficial in such varying situations.

## 6 | CONCLUSION

In conclusion, the current study provides further evidence that the I-I-Eye teaching aid is effective in increasing jurors' sensitivity to eyewitness issues. In contrast to other types of judicial instructions (e.g., Henderson instructions and Biggers criteria), the I-I-Eye can assist jurors to better distinguish between weak and strong eyewitness evidence. Evidence from other studies also seem to suggest that expert testimony can be improved with the complementary use of the I-I-Eye (Wise & Kehn, 2020). This finding is also important because several studies have shown that expert testimony for eyewitness identifications may create skepticism instead of sensitizing jurors to eyewitness evidence (e.g., Pezdek et al., 2010). If the broader pattern of results surrounding the I-I-Eye continues to be confirmed, it should provide an important foundation for the I-I-Eye to be regularly used in courtrooms with the aim of improving triers of fact evaluation of eyewitness evidence. This could lead to important applied implications, in terms of reducing miscarriages of justice that are based on contaminated eyewitness evidence. The current study also further contributes to this field by presenting an alternative version of the I-I-Eye teaching aid in a video format, which may prove to be an engaging and effective tool which can be used in trials involving eyewitness evidence. This short instructional video may help increase acceptance of a pre-trial educational aid in cases involving eyewitness testimony. The current study also shows that the I-I-Eye aids appear useful in a European as well as a U.S. legal system, for cases involving estimator as well as system variables.

## DATA AVAILABILITY STATEMENT

Data for Experiments 1 and 2 are available at: [https://osf.io/t46na/?view\\_only=ac69527fa8b04d10b9b374e1f01e3195](https://osf.io/t46na/?view_only=ac69527fa8b04d10b9b374e1f01e3195).

## ORCID

Renan Saraiva  <https://orcid.org/0000-0001-9461-7594>

## REFERENCES

- American Psychological Association (2012). Brief for amicus curiae, in support of petitioner. *Perry v. New Hampshire*, No. 10-8974, 565 U.S.
- Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology*, 20(1), 115–129.
- Bjørndal, L. D., McGill, L., Magnussen, S., Richardson, S., Saraiva, R., Stadel, M., & Brennen, T. (2020). Norwegian judges' knowledge of factors affecting eyewitness testimony: A 12-year follow-up. *Psychiatry, Psychology and Law*, 1–18.
- Bradfield, A. L., & Wells, G. L. (2000). The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria. *Law and Human Behavior*, 24(5), 581–594.
- 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.
- Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *Life Sciences Education*, 15(4), 1–6.
- Carmichael, M., Reid, A., & Karpicke, J. D. (2018). Assessing the impact of educational video on student engagement, critical thinking and learning. California: SAGE.
- Cherrett, T., Wills, G., Price, J., Maynard, S., & Dror, I. E. (2009). Making training more cognitively effective: Making videos interactive. *British Journal of Educational Technology*, 40(6), 1124–1134.
- Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, 42(6), 1777–1832.
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49, 219–243.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, 29(5), 575–604.
- Desmarais, S. L., & Read, J. D. (2011). After 30 years, what do we know about what jurors know? A meta-analytic review of lay knowledge regarding eyewitness factors. *Law and Human Behavior*, 35(3), 200–210.
- Dillon, M. K., Jones, A. M., Bergold, A. N., Hui, C. Y. T., & Penrod, S. D. (2017). Henderson instructions: Do they enhance evidence evaluation? *Journal of Forensic Psychology Research and Practice*, 17, 1–24.
- Douglas, A. B., & Jones, E. E. (2013). Confidence inflation in eyewitnesses: Seeing is not believing. *Legal and Criminological Psychology*, 18, 152–167.
- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law*, 19(1), 35–66.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109.
- Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535–549.
- Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). Doing meta-analysis with R: A hands-on guide.

- Innocence Project (2021). Eyewitness identification reform. Innocence Project. <https://innocenceproject.org/eyewitness-identification-reform/>
- Jones, A. M., Bergold, A. N., Dillon, M. K., & Penrod, S. D. (2017). Comparing the effectiveness of Henderson instructions and expert testimony: Which safeguard improves jurors' evaluations of eyewitness evidence? *Journal of Experimental Criminology*, 13, 29–52.
- Jones, A. M., Bergold, A. N., & Penrod, S. (2020). Improving juror sensitivity to specific eyewitness factors: Judicial instructions fail the test. *Psychiatry, Psychology and Law*, 27(3), 366–385.
- Jones, A. M., & Penrod, S. (2018). Improving the effectiveness of the Henderson instruction safeguard against unreliable eyewitness identification. *Psychology, Crime & Law*, 24(2), 177–193.
- Lim, J., Ko, H., Yang, J. W., Kim, S., Lee, S., Chun, M. S., & Park, J. (2019). Active learning through discussion: ICAP framework for education in health professions. *BMC Medical Education*, 19(1), 1–8.
- Magnussen, S., Eilertsen, D. E., Teigen, K. H., & Wessel, E. (2014). The probability of guilt in criminal cases: Are people aware of being 'beyond reasonable doubt'? *Applied Cognitive Psychology*, 28(2), 196–203.
- Magnussen, S., Melinder, A., Stridbeck, U., & Raja, A. Q. (2010). Beliefs about factors affecting the reliability of eyewitness testimony: A comparison of judges, jurors and the general public. *Applied Cognitive Psychology*, 24, 122–133.
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, 94(3), 339–354.
- Mueller-Johnson, K., Toglia, M. P., Sweeney, C. D., & Ceci, S. J. (2007). The perceived credibility of older adults as witnesses and its relation to ageism. *Behavioral Sciences & the Law*, 25(3), 355–375.
- National Research Council. (2014). *Identifying the culprit: Assessing eyewitness identification*. The National Academies Press.
- Neil v. Biggers. (1972). 409 U.S. 188.
- New Jersey v. Henderson, 27 A.3d 872 (2011).
- Nyman, T. J., Lampinen, J. M., Antfolk, J., Korkman, J., & Santtila, P. (2019). The distance threshold of reliable eyewitness identification. *Law and Human Behavior*, 43, 527–541.
- Papailiou, A. P., Yokum, D. V., & Robertson, C. T. (2015). The novel New Jersey eyewitness instruction induces skepticism but not sensitivity. *PLoS One*, 10, 1–16.
- Pawlenko, N. B., Safer, M. A., Wise, R. A., & Holfeld, B. (2013). A teaching aid for improving jurors' assessments of eyewitness accuracy. *Applied Cognitive Psychology*, 27, 190–197.
- Pezdek, K., Avila-Mora, E., & Sperry, K. (2010). Does trial presentation medium matter in jury simulation research? Evaluating the effectiveness of eyewitness expert testimony. *Applied Cognitive Psychology*, 24(5), 673–690.
- Police and Criminal Evidence Act (2017). Code of practice for the identification of persons by police officers. Norwich: TSO (The Stationery Office)
- Ramirez, G., Zemba, D., & Geiselman, R. E. (1996). Judges' cautionary instructions on eyewitness testimony. *American Journal of Forensic Psychology*, 14, 31–66.
- Roebbers, C. M., & Schneider, W. (2000). The impact of misleading questions on eyewitness memory in children and adults. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 14(6), 509–526.
- Safer, M. A., Murphy, R. P., Wise, R. A., Bussey, L., Millett, C., & Holfeld, B. (2016). Educating jurors about eyewitness testimony in criminal cases with circumstantial and forensic evidence. *International Journal of Law and Psychiatry*, 47, 86–92.
- Sauer, J., 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(4), 337–347.
- Schwarzer, G. (2007). Meta: An R package for meta-analysis. *R News*, 7(3), 40–45.
- Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education*, 165, 104132.
- Technical Working Group for Eyewitness Evidence. (2014). *Eyewitness evidence: A guide for law enforcement*. National Institute of Justice.
- Vyond (2019). Computer Software. <https://www.vyond.com/>.
- Wells, L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 22, 603–647.
- Wiggins, B. L., Eddy, S. L., Grunspan, D. Z., & Crowe, A. J. (2017). The ICAP active learning framework predicts the learning gains observed in intensely active classroom experiences. *AERA Open*, 3(2), 23328584 17708567.
- Wise, R. A., Fishman, C. S., & Safer, M. A. (2009). How to analyze the accuracy of eyewitness testimony in a criminal case. *Connecticut Law Review*, 42, 435–513.
- Wise, R. A., & Kehn, A. (2020). Can the effectiveness of eyewitness expert testimony be improved? *Psychiatry, Psychology and Law*, 27(2), 315–330.
- Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science*, 13(3), 324–335.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Saraiva, R., Bertoldo, G., Bjørndal, L. D., Bunghez, C., Lofthus, I. S., McGill, L., Richardson, S., & Stadel, M. (2022). Improving the evaluation of eyewitness evidence in legal decision-making: Testing an active versus passive teaching aid. *Applied Cognitive Psychology*, 36(3), 520–535. <https://doi.org/10.1002/acp.3938>