

## The Effects of Verbal Descriptions on Eyewitness Memory: Implications for the Real-World<sup>☆</sup>



Laura Mickes\*

University of London, United Kingdom

The criminal justice system depends on verbal accounts of crimes. Can the act of reporting a crime harm eyewitness memory for the perpetrator of that crime? The answer is yes according to the *verbal overshadowing effect*. The verbal overshadowing effect describes the finding that memory is adversely affected after verbally describing a previously presented item (e.g., face). Often in studies of the verbal overshadowing effect, participants watch a video of a mock crime, describe the perpetrator (verbal condition) or engage in another task (control condition). In many of these studies, including the original (Schooler & Engstler-Schooler, 1990) and replication studies (Alogna et al., 2014), memory for a perpetrator is tested on target-present lineups, and, if described, the perpetrator is less often identified. However, it is unknown whether or not the lower identification rate is due to reduced discriminability or due to more conservative responding after providing a description. The verbal overshadowing effect ought to be defined as a reduction in discriminability, which is measured by taking both the correct ID rates (from target-present lineups) and false ID rates (from target-absent lineups) into consideration. Another important and independent measure is the reliability of identifications (i.e., the positive predictive value of a suspect identification made with a given level of confidence). As matters stand, the take-home message is this: too little information currently exists to allow for an assessment of the effects of verbal descriptions on discriminability and reliability; thus, the field is not yet in a position to offer clear guidance for practice in the criminal justice system.

**Keywords:** Eyewitness memory, Verbal overshadowing, Discriminability, Reliability, Confidence–accuracy relationship, Policy recommendations

### Reporting Crimes and Making Identifications

From a criminal offence to completion of the ensuing court case, the criminal justice system follows a linear process. The entire process usually takes at least several months, and as shown in Figure 1 may include the crime, report, investigation (if deemed worthy by the police), eyewitness identification (ID) procedure administration, formal charge against the suspect, and court case. The timescale in Figure 1 represents averages of indicted cases in the UK (UK Ministry of Justice, 2011). Reporting a crime to the authorities inevitably involves describing details of the crime and the perpetrator(s). Emergency services, call dispatchers, and investigating officers are trained to ask questions about the crime in such a way that as much accurate

information as possible is gathered in a non-suggestive way (Technical Working Group for Eyewitness Evidence, 1999), and online self-report forms follow a similar structure (College of Policing, 2013). To answer questions about the perpetrator, eyewitnesses are asked to describe the individual. If needed, eyewitnesses are prompted to consider the perpetrator's age, gender, ethnicity, height, build, distinguishing characteristics, etc. (Association of Chief Police Officers, 2016). If police later identify a suspect, as part of the investigation, a lineup procedure may be administered to eyewitnesses.

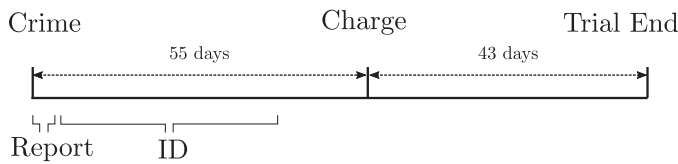
A lineup consists of the police suspect (who may or may not be the perpetrator) and several other individuals who physically resemble the perpetrator, called “fillers.” The lineup members

#### Author Note

This work was supported by the Economic and Social Research Council [ES/L012642/1]. The content is solely the responsibility of the author and does not necessarily reflect the views of the Economic and Social Research Council.

<sup>☆</sup> Please note that this paper was handled by the current editorial team of JARMAC.

\* Correspondence concerning this article should be addressed to Laura Mickes, Department of Psychology, Royal Holloway, University of London, Egham TW20 0EX, United Kingdom. Contact: [laura.mickes@rhul.ac.uk](mailto:laura.mickes@rhul.ac.uk)



**Figure 1.** Criminal justice system case progression in the UK.

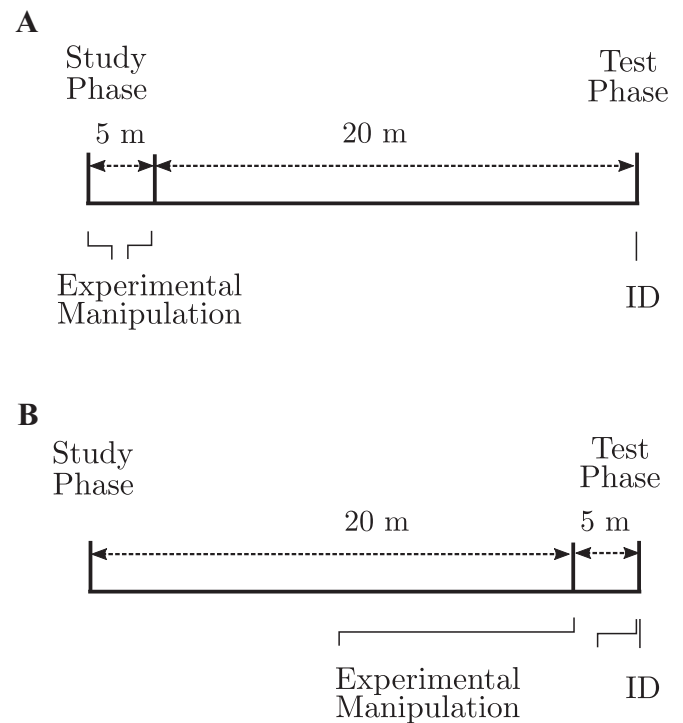
are all presented via photos or videos, and the witness attempts to identify the perpetrator (ID in [Figure 1](#)). What if the task of verbally describing the perpetrator has a detrimental effect on memory for that very perpetrator?

### The Verbal Overshadowing Effect

That is the implication of a finding first reported nearly 30 years ago ([Schooler & Engstler-Schooler, 1990](#)). In a set of experiments, participants viewed a video of a mock robbery during the study phase, and either described the perpetrator (verbal condition) or engaged in a control task (control condition). Memory for the perpetrator, or target, was tested on an 8-person simultaneous target-present lineup. Surprisingly, participants in the verbal condition were less able to correctly identify the target than those who were not asked to verbally describe the perpetrator. This counterintuitive finding, termed the *verbal overshadowing effect*, inspired much followup research with mixed results (e.g., [Dodson, Johnson, & Schooler, 1997](#); [Finger, 2002](#); [Finger & Pezdek, 1999](#); [Kitagami, Sato, & Yoshikawa, 2002](#); [Nakabayashi, Lloyd-Jones, Butcher, & Liu, 2012](#); [Smith & Flowe, 2014](#); [Wickham & Swift, 2006](#)). Because of this, and because a meta-analysis revealed a much smaller effect than the original experiments ([Meissner & Brigham, 2001](#)), two of the original experiments were the object of a large direct replication effort ([Alogna et al., 2014](#)).

[Figure 2](#) shows a schematic of the experimental design of the two replication experiments. In both experiments, the procedure was delineated by the study phase (presentation of the mock crime video) and the test phase (memory tested on an 8-person lineup). The only difference between the experiments was the timing of the experimental manipulation (where participants either verbally described the perpetrator or did not). Clearly, the experimental analog is a much shorter version of the protracted criminal justice system in [Figure 1](#), which is a point discussed later. In Experiment 1, the experimental manipulation occurred immediately after the study phase ([Figure 2A](#)) and in Experiment 2, the experimental manipulation occurred 20 min after the study phase ([Figure 2B](#)). The effect replicated. In both experiments, the correct ID rate (i.e., the proportion of guilty suspects identified from target-present lineups) was lower in the verbal condition, but markedly lower when the verbal description was given 20 min after the study phase and immediately before the test (and the effect sizes were small, especially in Experiment 1).

However, by comparing only correct ID rates, it is unclear whether the difference is due to a difference in discriminability (the ability to distinguish innocent from guilty suspects) or response bias (the likelihood of choosing a lineup member)



**Figure 2.** Procedural order of the replication studies for Experiment 1 (A) and Experiment 2 (B) in [Alogna et al. \(2014\)](#).

([Clare & Lewandowsky, 2004](#); [Meissner & Brigham, 2001](#)). To disentangle the two possible explanations for the difference, it is necessary to include target-absent lineups in the experimental design. By doing so, false ID rates (i.e., the proportion of innocent suspects identified from target-absent lineups) can be taken into account and discriminability can be measured separately from response bias ([Mickes & Wixted, 2015](#)).

### Discriminability in Verbal Overshadowing: A Matter of Concern for Policymakers

A veridical verbal overshadowing effect ought to be defined by a reduction in discriminability (i.e., lower correct ID rates and higher false ID rates) in the verbal condition compared to the control condition. Discriminability cannot be measured by only a reduction in correct ID rates. It follows that the results of the replication studies cannot inform whether or not discriminability is affected after providing a verbal account ([Mickes & Wixted, 2015](#); [Rotello, Heit, & Dube, 2015](#)). To be informed about discriminability, receiver operating characteristic (ROC) analysis, which measures objective discriminability of lineup data, needs to be conducted ([Gronlund, Wixted, & Mickes, 2014](#); [National Research Council, 2014](#); [Wixted & Mickes, 2012](#)).

ROC analysis was recently introduced to measure discriminability in lineup data ([Wixted & Mickes, 2012](#)), and there is currently some resistance to its use in the field of eyewitness identification research ([Wells, Smalarz, & Smith, 2015](#); [Wixted & Mickes, 2015a, 2015b](#)). Some researchers continue to support the use of the diagnosticity ratio (DR; correct ID rate/false ID rate) to measure discriminability in preference to ROC analysis, arguing that ROC analysis is not appropriate for lineups ([Wells](#)

**Table 1**

Correct ID Rates from *Alogna et al. (2014)*, Hypothetical False ID Rates, and Corresponding  $d'$  Scores and Diagnosticity Ratios (DR). The Hypothetical Values are in Bold Font

	Different discriminability		Equal discriminability	
	Verbal	Control	Verbal	Control
Experiment 1				
Correct ID rate	0.51	0.55	0.51	0.55
False ID rate	<b>0.11</b>	<b>0.02</b>	<b>0.02</b>	<b>0.02</b>
$d'$	<b>1.27</b>	<b>2.18</b>	<b>2.18</b>	<b>2.18</b>
DR	<b>4.81</b>	<b>27.49</b>	<b>32.48</b>	<b>27.50</b>
Experiment 2				
Correct ID rate	0.38	0.54	0.38	0.54
False ID rate	<b>0.06</b>	<b>0.02</b>	<b>0.01</b>	<b>0.02</b>
$d'$	<b>1.26</b>	<b>2.16</b>	<b>2.15</b>	<b>2.15</b>
DR	<b>6.37</b>	<b>27.18</b>	<b>54.29</b>	<b>27.00</b>

et al., 2015). However, it is not clear how one can successfully argue that it is acceptable to measure overall correct and false ID rates from lineups, which are needed to compute the DR, while at the same time arguing that it is unacceptable to compute all other correct and false ID rates, which are needed to plot the ROC (e.g., by setting a more conservative standard and not counting any ID made with very low confidence). Moreover, because the DR confounds response bias and discriminability (Gronlund et al., 2014), it is not the pure measure of discriminability that ROC analysis is.

ROC analysis is widely accepted as the preferred measure of discriminability in other fields (e.g., diagnostic medicine, experimental psychology, machine learning, physics, etc.), and it was recently deemed superior to the DR by a prestigious committee of the National Academy of Sciences charged with evaluating research methodologies and empirical findings in eyewitness identification (National Research Council, 2014). Given its widespread use in other fields and its recent backing by the National Research Council, my own view is that it is only a matter of time before eyewitness identification researchers as a whole accept ROC analysis as the proper way to measure discriminability. Nevertheless, for now, it also seems fair to say that others disagree with my position on this issue. Only time will tell how this debate will ultimately be resolved.

To conduct ROC analysis, correct ID rates are plotted against false ID rates, resulting in ROC curves for each condition. The larger the area under the ROC curve, the better the discriminability. In other words, the larger the area under the ROC curve, the better the identification procedure is at distinguishing between innocent and guilty suspects. Thus, for there to be a verbal overshadowing effect, the area under the verbal condition ROC curve would need to be smaller than that of the control condition ROC curve.

Though false ID rates, essential for the construction of ROC curves, are not available from the replication studies (because target-absent lineups were not included), hypothetical false ID rates can be used to demonstrate the point about discriminability. Table 1 shows the correct ID rates reported in *Alogna et al. (2014)* and hypothetical false ID rates. The

**Table 2**

Concerns, Relevant Analyses, and Goals for Different Decision-Makers

	Policymakers	Courts
Concern	Discriminability	Reliability
Analysis	Receiver operating characteristic	Confidence-accuracy characteristic
Goal	High discriminability	IDs made with high confidence are highly accurate

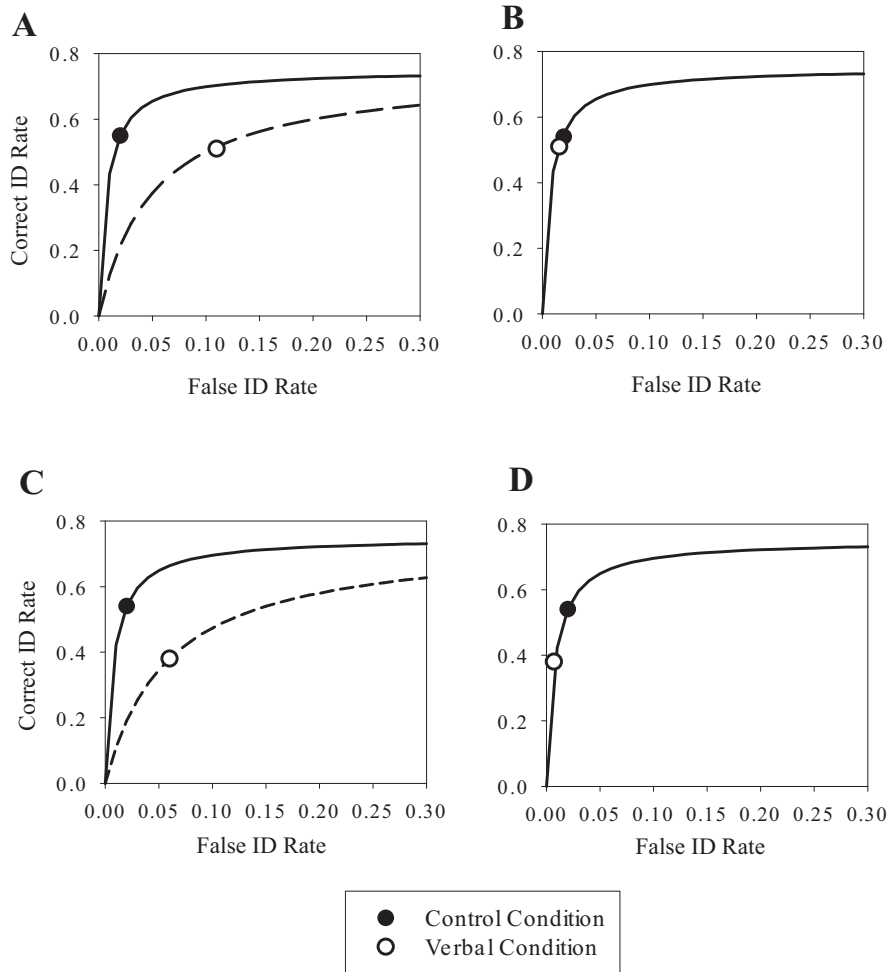
hypothetical false ID rates are in bold font. A parametric measure of discriminability,  $d'$ , and diagnosticity ratio, DR, which are computed using the correct and false ID rates are also shown in bold font. Figure 3 shows four possible ROC outcomes of the replication Experiment 1 (A and B) and Experiment 2 (C and D). In both experiments, it is possible, given the available data, that the verbal condition falls on a lower ROC, as shown in Figure 3A and C. If the data yielded this pattern, then that would be a clear difference in discriminability, and one could then conclude that there is a verbal overshadowing effect.

It is also possible, given the available data, that the verbal condition falls on the same ROC as the control condition, as shown in Figure 3B and D. If the data yielded this pattern, then there would not be a difference in discriminability, and thus no verbal overshadowing effect, but there would be a difference in response bias between the two conditions. Though this point can be demonstrated by using overall correct and false ID rates, ideally confidence ratings associated with the identifications would be used in the analysis so that the entire locus of ROC operating points per condition can be plotted (Mickes, Flowe, & Wixted, 2012).

When and if differences in discriminability arise after a verbal description is one fact worth knowing, and the ROC results should be used to aid decision-makers (e.g., police chiefs) charged with making procedural endorsements. Although ROC analysis provides essential information to policymakers, it does not provide particularly useful information to triers of fact (judges and juries). Triers of fact are interested in the reliability of an ID, and that information is provided by an altogether different analysis. The different concerns of different decision-makers and the relevant analyses are shown in Table 2.

### Reliability in Verbal Overshadowing: A Matter of Concern for the Courts

Whether or not discriminability is affected by verbal reports is not a matter for judges and jurors. In the court of law, the concern is about reliability. Reliability is measured by the positive predictive value (PPV) of suspect identifications. The PPV is the probability that a suspect who was identified by a witness is the perpetrator. For the discussion that follows, equal base rates (i.e., half target-present lineups and half target-absent lineups) are assumed for the sake of simplicity. Generally speaking, PPV varies directly with the base rate of target present lineups. The PPV can be measured in several different ways. One way is to compute a DR. Using the data in Table 1 and the corresponding ROC curves in Figure 3B to compute DR values,



**Figure 3.** ROC curves constructed with correct ID rates and hypothetical false ID rates for Experiment 1 (A and B) and Experiment 2 (C and D) of [Alogna et al. \(2014\)](#). The ROC curves in A and C show a clear discriminability advantage for the control condition. The ROC curves in B and D show no discriminability difference, but more conservative responding for the verbal condition. With the existing data, it is possible to have different or equal discriminability, and which it is remains unknown.

for the verbal condition  $DR = 32.5$  and for the control condition  $DR = 27.5$ . The higher DR scores in the verbal condition would mean that reliability is better in that condition (despite having the same discriminability). A related but more complete way to measure reliability is to conduct confidence-accuracy characteristic (CAC) analysis ([Mickes, 2015](#)). Like the DR, CAC analysis uses only correct and incorrect suspect identifications. Thus, the dependent measure, PPV, is defined by

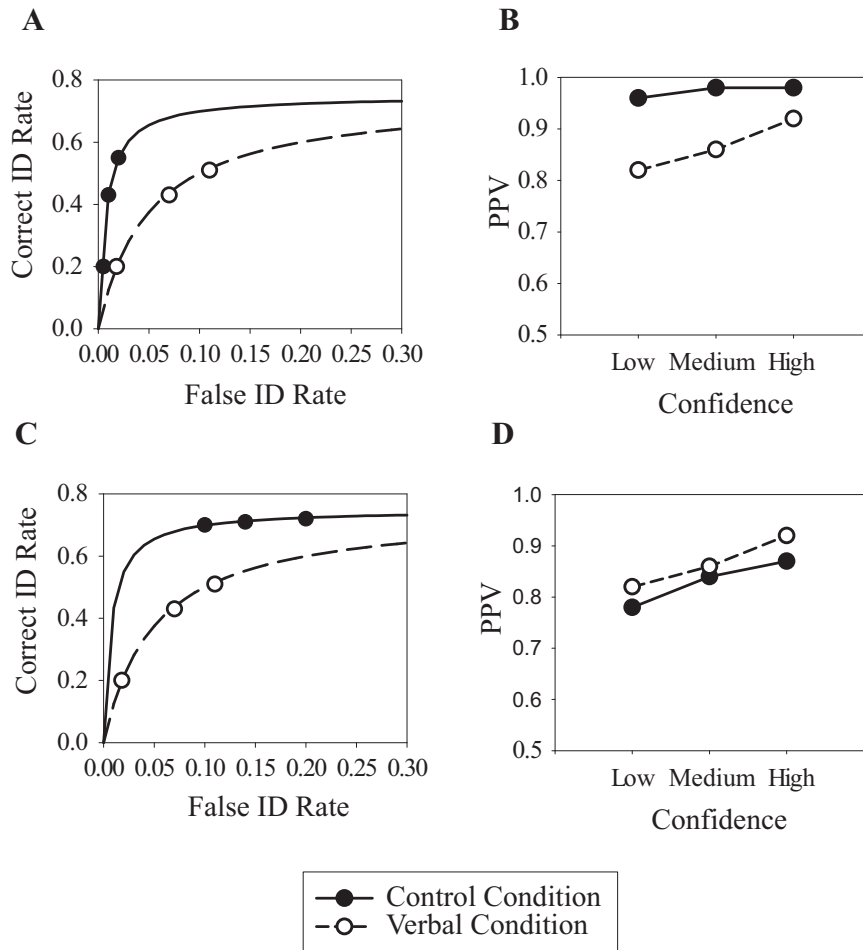
$$PPV = \frac{S_g}{S_g + S_i}, \quad (1)$$

where  $S_g$  is the number of correct IDs (i.e., the number of guilty suspects identified from target-present lineups).  $S_i$  refers to the number of innocent suspects identified from target-absent lineups. When the base rates are equal,  $S_g$  and  $S_i$  can be thought of as the correct and false ID rates, respectively.

Using the indicative ROC curves from [Figure 3A](#), three ROC points can be computed for low, medium, and high confidence levels for both control and verbal conditions, shown in [Figure 4A](#). The PPV is computed for every level of confidence and shown in [Figure 4B](#). In [Figure 4A](#), the ROC is lower for

the verbal condition; in [Figure 4B](#), the PPV is also lower for the verbal condition and by the same amount. In this case the verbal condition has simultaneously lower discriminability ([Figure 4A](#)) and lower reliability ([Figure 4B](#)). Generally higher discriminability is associated with higher reliability, but this is not always the case. This can be seen clearly from [Figure 4C and D](#), where the same ROC curves are used but the ROC points are shifted towards more liberal responding for the control condition. Discriminability is still lower for the verbal condition ([Figure 4C](#)), but the PPV is actually higher in the verbal condition ([Figure 4D](#)).

The point is that the two analyses answer different questions and, while they likely agree most of the time in practice, they are potentially dissociable (as they are in [Figure 4C and D](#)). Because of this possibility, it is necessary to analyze the data using both ROC and CAC analyses. Furthermore, the fact that the results of CAC analysis are easier to understand give it a distinct advantage over the DR. For example, 98% versus 92% accurate makes more sense than DR values of 40 versus 11. Therefore, research shedding light on the reliability of identifications using CAC analysis is needed. Moreover, when and if differences in



**Figure 4.** Hypothetical ROC curves and CAC curves. Both ROC curves (A and C) show lower discriminability for the verbal condition, but the associated PPVs differ with higher PPV in the control condition (B), and higher PPV in the verbal condition (D). This illustrates why ROC analysis and CAC analysis need to be conducted to measure discriminability and reliability, respectively.

reliability arise after a verbal description is worth knowing, and the results should be used to aid decision-makers (e.g., judges and jurors) charged with making judgments about culpability (see Table 2).

### Conclusions

Collecting verbal reports from eyewitnesses of crimes is an inescapable necessity. As the research currently stands, we lack sufficient information about discriminability (a matter of concern for policymakers) and reliability (a matter of concern for judges and jurors) to make recommendations to guide practice. Furthermore, different patterns may arise that are contingent on the intervals between exposure and description and identification. After all, different retention intervals differentially affect correct ID rates (Alogna et al., 2014). Does that finding reflect a difference in discriminability or a difference in response bias?

There was a glimpse in the replication studies that discriminability is lower after time lapses between the crime and the report. In Experiment 1, in which the experimental manipulation occurred immediately after the study phase and 20 min before the identification test (see Figure 2A), the correct ID rates and the filler ID rates (i.e., identifications made to fillers

in the target-present lineup) were lower in the verbal condition than in the control condition. This pattern suggests a shift in response bias. However, in Experiment 2, in which the experimental manipulation occurred 20 min after the study phase and immediately before the lineup test (see Figure 2B), the filler ID rates were no different for the verbal condition versus the control condition despite the fact that correct ID rates were lower in the former condition. This pattern suggests that there was not a shift in response bias, but lower discriminability. However, ROC analysis is required to definitively answer the question of discriminability and importantly, if the ROC curves are repeatedly lower with longer delays between crime and description, then the police should encourage eyewitnesses to report crimes as soon as possible.

What about reliability? That is what CAC analysis will reveal. If the CAC curves are higher, but the ROC curve is lower in the verbal condition (a possibility demonstrated in Figure 4C and D), then the conclusion would be that the verbal overshadowing effect is real, but, compared to the control condition, identifications made with high confidence from the verbal condition are more reliable anyway. This scenario may arise because the task of verbally describing the perpetrator is a challenging one. That fact may be appreciated by the participants in that condition



who in turn may be more cautious to make an identification with high confidence (Clare & Lewandowsky, 2004). Both hypothetical CAC scenarios in Figure 4 are possible. Another important question is whether the timing of the verbal description and later identification differentially affect reliability (as it might with discriminability). The research has yet to, but needs to, be conducted.

The effects of verbal descriptions on eyewitness memory is worth investigating using procedures that are more protracted to mimic the experience of real eyewitnesses (i.e., eyewitnesses do not provide a verbal description immediately after seeing the perpetrator, nor would a lineup procedure be administered immediately after describing the perpetrator; Mickes & Wixted, 2015). In addition to extending the procedural timeline, target-absent lineups need to be included, and the appropriate analyses need to be conducted. Confidence should be collected (1) to measure discriminability (with ROC analysis) and to measure reliability (with CAC analysis), and (2) because confidence at first identification is diagnostic of accuracy (e.g., Brewer & Wells, 2006; Horry, Palmer, & Brewer, 2012; Palmer, Brewer, Weber, & Nagesh, 2013; Sauer, Brewer, Zweck, & Weber, 2009; Wixted, Mickes, Clark, Gronlund, & Roediger, 2015). Once we have a body of work that replicates and researchers have come to a general consensus about the interpretations of the results, then we can guide practice. But we are not quite there yet.

## References

- Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, S., Birch, S., & Birt, A. R. (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science*, 9, 556–579. <http://dx.doi.org/10.1177/1745691614545653>
- Association of Chief Police Officers. (2016). *True vision association of chief police officers*. <http://report-it.org.uk/home>
- Brewer, N., & Wells, G. L. (2006). The confidence–accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, 12, 11–30.
- Clare, J., & Lewandowsky, S. (2004). Verbalizing facial memory: Criterion effects in verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 739–755. <http://dx.doi.org/10.1037/0278-7393.30.4.739>
- College of Policing. (2013). *Investigation: Investigative interviewing*. <https://www.app.college.police.uk/app-content/investigations/investigative-interviewing/#structuring-a-witness-interview>
- Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, 25, 129–139. <http://dx.doi.org/10.3758/BF03201107>
- Finger, K. (2002). Mazes and music: Using perceptual processing to release verbal overshadowing. *Applied Cognitive Psychology*, 16, 887–896. <http://dx.doi.org/10.1002/acp.922>
- Finger, K., & Pezdek, K. (1999). The effect of the cognitive interview on face identification accuracy: Release from verbal overshadowing. *Journal of Applied Psychology*, 84, 340–348.
- Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using receiver operating characteristic analysis. *Current Directions in Psychological Science*, 23, 3–10. <http://dx.doi.org/10.1177/0963721413498891>
- Horry, R., Palmer, M. A., & Brewer, N. (2012). Backloading in the sequential lineup prevents within-lineup criterion shifts that undermine eyewitness identification performance. *Journal of Experimental Psychology: Applied*, 18, 346–360.
- Kitagami, S., Sato, W., & Yoshikawa, S. (2002). The influence of test-set similarity in verbal overshadowing. *Applied Cognitive Psychology*, 16, 963–972. <http://dx.doi.org/10.1002/acp.917>
- Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology*, 15, 603–616. <http://dx.doi.org/10.1002/acp.728>
- Mickes, L. (2015). Receiver operating characteristic analysis and confidence accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, 4, 93–102. <http://dx.doi.org/10.1016/j.jarmac.2015.01.003>
- Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied*, 18, 361–376.
- Mickes, L., & Wixted, J. T. (2015). On the applied implications of the “verbal overshadowing effect”. *Perspectives on Psychological Science*, 10, 400–403. <http://dx.doi.org/10.1177/1745691615576762>
- Nakabayashi, K., Lloyd-Jones, T. J., Butcher, N., & Liu, C. H. (2012). Independent influences of verbalization and race on the configural and featural processing of faces: A behavioral and eye movement study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 61–77. <http://dx.doi.org/10.1037/a0024853>
- National Research Council. (2014). *Identifying the culprit: Assessing eyewitness identification*. Washington, DC: The National Academies Press. <http://dx.doi.org/10.17226/18891>
- Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence–accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, 19, 55–71.
- Rotello, C. M., Heit, E., & Dube, C. (2015). When more data steer us wrong: replications with the wrong dependent measure perpetuate erroneous conclusions. *Psychonomic Bulletin and Review*, 22, 944–954. <http://dx.doi.org/10.3758/s13423-014-0759-2>
- Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2009). The effect of retention interval on the confidence–accuracy relationship for eyewitness identification. *Law and Human Behavior*, 34, 337–347. <http://dx.doi.org/10.1007/s10979-009-9192-x>
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36–71. [http://dx.doi.org/10.1016/0010-0285\(90\)90003-M](http://dx.doi.org/10.1016/0010-0285(90)90003-M)
- Smith, H. M. J., & Flowe, H. D. (2014). Roc analysis of the verbal overshadowing effect: Testing the effect of verbalisation on memory sensitivity. *Applied Cognitive Psychology*, 29, 159–168. <http://dx.doi.org/10.1002/acp.3096>
- Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: United States Department of Justice, Office of Justice Programs, National Institute of Justice. <https://www.ncjrs.gov/pdffiles1/nij/178240.pdf>
- UK Ministry of Justice. (2011, June). *Time intervals survey of criminal proceedings in magistrates’ courts*. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/217768/tis-bulletin-0611.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/217768/tis-bulletin-0611.pdf)
- Wells, G. L., Smalarz, L., & Smith, A. M. (2015). Roc analysis of lineups does not measure underlying discriminability and has limited value. *Journal of Applied Research in Memory and Cognition*, 4, 313–317. <http://dx.doi.org/10.1016/j.jarmac.2015.08.008>

- Wickham, L. H. V., & Swift, H. (2006). Articulatory suppression attenuates the verbal overshadowing effect: A role for verbal encoding in face identification. *Applied Cognitive Psychology, 20*, 157–169. <http://dx.doi.org/10.1002/acp.1176>
- Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives on Psychological Science, 7*, 275–278. <http://dx.doi.org/10.1177/1745691612442906> <http://pps.sagepub.com/content/7/3/275.abstract>
- Wixted, J. T., & Mickes, L. (2015a). Evaluating eyewitness identification procedures: Roc analysis and its misconceptions. *Journal of Applied Research in Memory and Cognition, 4*, 318–323. <http://dx.doi.org/10.1016/j.jarmac.2015.08.009>
- Wixted, J. T., & Mickes, L. (2015b). Roc analysis measures objective discriminability for any eyewitness identification procedure. *Journal of Applied Research in Memory and Cognition, 4*, 329–334. <http://dx.doi.org/10.1016/j.jarmac.2015.08.007>
- Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L., III. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist, 70*, 515–526.

Received 1 May 2016;  
received in revised form 3 July 2016;  
accepted 3 July 2016