On the applied implications of the "Verbal Overshadowing Effect"

Laura Mickes[1] & John T. Wixted[2]

[1]Royal Holloway, University of London, [2]University of California, San Diego

Abstract

Schooler and Engstler-Schooler (1990) found that participants who wrote out a description of the perpetrator's face after watching a simulated crime video were subsequently less likely to identify that perpetrator from a photo lineup compared to participants in a control condition (i.e., the correct ID rate was reduced). The first registered replication report confirmed this "verbal overshadowing effect" (Alogna et al, 2014). Does this result indicate a reduced ability to recognize the person who was verbally described, or does it instead reflect more conservative responding? The answer depends on the still unknown likelihood of identifying an innocent suspect from a lineup (the false ID rate). Assuming the reduced correct ID rate does reflect memory impairment, should the legal system be advised to give less weight to a suspect identification if the witness previously provided a verbal description of the perpetrator? Intuitively, the answer is "yes," but without knowing the false ID rate, it is unclear if a suspect identification following a verbal description should be given less weight *or more* weight. This is true even if the correct and false ID rates show that verbal descriptions impair memory. In our view, psychologists should withhold giving advice to the legal system about the effect of verbal descriptions on suspect identifications until the issue is investigated by including lineups that contain an innocent suspect.

In a well-known study investigating the effect of verbally describing a face on subsequent memory for that same face, Schooler and Engstler-Schooler (1990) asked participants to watch a video of a simulated bank robbery and later tested their ability to identify the robber from a 6-person lineup. They found that participants who wrote out a description of the perpetrator's face after watching the video were subsequently less likely to identify the robber from the lineup compared to participants in a control condition who, instead of describing the perpetrator's face, generated a list of capital cities. In part because this finding could have ramifications for police practices, *Perspectives on Psychological Science* selected this study for its first registered replication report (RRR). Two variants of the study were run; they differed only in the order of the participant's tasks. In RRR1, participants watched the video, then immediately described the face, then engaged in a 20-minute distractor task, whereas in RRR2, participants engaged in the distractor task before describing the face. Upon completing these activities, the participants attempted to identify the perpetrator from the photo lineup. The main dependent measure was the proportion of lineups from which witnesses correctly identified the perpetrator (the correct ID rate).

For RRR1, the meta-analytic effect across 31 replications (correct ID rate in the verbal description minus the correct ID rate in the control condition) was -4.01% (95% confidence interval: -7.15% to -0.87%). For RRR2, the meta-analysis of 22 studies showed a difference of -16.31% (95% confidence interval: -20.47% to -12.14%). Thus, the decrement was statistically significant either way but was substantially larger when the verbal description occurred 20 minutes after the video and just before the lineup test[1].

In describing the potential policy implications of the replication effort, Alogna et al. (2014) wrote "If asking a witness to verbally describe the person they saw

substantially impairs their ability to recognize that person later, then eyewitness identification should be weighted less if the witness had provided a description earlier" (p. 557). However, that policy does not automatically follow from the antecedent. Even if verbal descriptions impair one's ability to later recognize a person, it might be the case that *more weight* should be attached to an identification made by an eyewitness who provided a description of the perpetrator. In this comment, we explain why.

Two Possible Explanations of the Verbal Shadowing Effect

To appreciate why the existence of the verbal overshadowing effect does not necessarily imply that identifications made by eyewitnesses who provide a verbal description should be discounted, it is important to consider the fact that in both the original study and the 31 replications only target-present lineups were used (i.e., lineups that contained the guilty suspect). The correct ID rate is also known as the hit rate. Both the original study and its replications show that the hit rate is significantly lower when the perpetrator's face is described compared to when it is not described. That difference in the hit rates across the two conditions *is* the verbal overshadowing effect. As has been noted before, and as was also noted in the registered replication report (see page 570 of Algona et al., 2014), a decrease in the hit rate can occur either because memory has been impaired (i.e., discriminability – the ability to distinguish what was seen from what was not seen – has been reduced) or because responding has become more conservative (i.e., inclination to choose has been reduced – e.g., Clare & Lewandowsky, 2004; Meissner & Brigham, 2001). To find out which explanation applies when memory is tested using a lineup procedure, the corresponding false alarm rates for the two conditions must be determined as well[2]. To obtain that information, additional participants would need to be tested using target-absent

lineups (i.e., lineups that contain an innocent suspect instead of the guilty suspect).

The proportion of target-absent lineups from which the innocent suspect is incorrectly

identified is the false ID rate (i.e., the false alarm rate). Instead of choosing the

suspect, eyewitnesses presented with target-present or target-absent lineups may

choose a filler or decide that the perpetrator is not in the lineup. Our focus is on

suspect IDs because only those IDs contribute to correct and wrongful convictions.

Conservative Responding

There have been studies of the verbal overshadowing effect that have included

target-absent lineups (Clare & Lewandowsky, 2004; Meissner, 2002, Memon & Rose,

2002, Sauerland, Holub, & Sporer, 2008; Dehon, Vanootighem, & Brédart, 2013).

One study concluded that the effect arose because participants in the verbal

description condition were more conservative than participants in the control

condition (Clare & Lewandowsky, 2004). They found that participants were more

reluctant to make an ID from any lineup (target-present or target-absent). As a result,

fewer guilty suspects were identified after verbally describing the perpetrator's face

(that is the verbal overshadowing effect), but fewer innocent suspects were identified

as well. In other words, both the correct ID and the false ID rate were lower in the

verbal overshadowing condition (see Table 1).

Insert Table 1 About Here

As noted in a number of recent articles, a key consideration is that the

probative value of an ID (in other words, the trustworthiness of an ID) increases as

responding becomes more conservative (e.g., Wixted & Mickes, 2012). This

phenomenon is invariably observed in receiver operating characteristic (ROC) data

and is predicted by the standard signal detection model of recognition memory performance (Wixted & Mickes, 2014). Thus, if inducing more conservative responding were the *only* effect of providing a verbal description (i.e., if that is why the correct ID rate decreases), then the identifications made by the witnesses in the verbal condition would be *more trustworthy* (not less trustworthy) than the identifications made by the witnesses in the control condition. Indeed, this very phenomenon is evident in the data reported by Clare and Lewandowsky (2004). In their Experiment 1, the correct and false ID rates computed separately from target-present and target-absent lineups in the control condition were .80 and .13, respectively (see Table 1). The probative value of a suspect ID in this condition is given by the diagnosticity ratio, .80 / .13 = 6.15. The correct and false ID rates from the verbal description condition were .63 and .08, respectively (.63 / .08 = 7.88). Thus, despite the fact that the correct ID rate decreased when the suspect was verbally described, the probative value of a suspect ID – that is, the trustworthiness of a suspect ID – increased.

Reduced Discriminability

Having both correct and false ID rates available allows one to compute the trustworthiness of an ID made from the different experimental conditions (information that is necessary to determine policy implications of the verbal overshadowing effect), but those values alone do not indicate if verbal descriptions impair discriminability. In this context, discriminability refers to the ability to distinguish between the face that was seen in the video from faces that were not (including the face of an innocent suspect). Although one could compute d' to measure discriminability (see Mickes, Moreland, Clark, & Wixted, 2014), a better approach would be to either perform ROC analysis or use a forced-choice procedure.

As described in classic signal-detection texts, these methods can more definitively indicate whether or not verbal descriptions influence discriminability (Green & Swets, 1966; Macmillan & Creelman, 2005). Knowing whether or not verbal descriptions affect discriminability is essential for theory development. At the moment, it is not clear if the relevant theory should address the effects of verbal descriptions on discriminability, response bias, or both.

Policy Implications

If the verbal overshadowing effect is determined to reflect reduced discriminability, would it mean that an ID made following a verbal description is less trustworthy than it otherwise would be? Not necessarily. Even in that case, verbal descriptions might *also* induce sufficiently conservative responding that the probative value of an ID would still increase. In fact, this exact state of affairs may apply to another line of research that has compared simultaneous vs. sequential lineups.[3]

With the forgoing considerations in mind, imagine that suspect IDs made in the verbal overshadowing condition were found to have lower probative value (i.e., lower accuracy) compared to IDs made in the control condition. This would occur, for example, if verbal descriptions impair discriminability (i.e., yield a lower ROC) without also inducing a conservative response bias. Under those conditions, would it finally be safe for psychologists to advise the legal system to attach less weight to IDs made by witnesses who provided a description of the perpetrator's face? It might be. We say "might" because, even here, there are additional factors to take into consideration. For example, it could be argued that, rather than discounting suspect IDs made by witnesses who provided a verbal description, a better approach might be to have the lineup administrator induce more conservative responding in such witnesses before they make an ID (e.g., by encouraging them not to make an ID

unless they are confident of being correct). This would have the effect of increasing the probative value of their suspect IDs, thereby offsetting the negative effect of providing a verbal description. At a minimum, policymakers should be advised that both options are available, and they could decide which approach to use.

In our view, experimental psychologists should offer no advice to the legal system based on the results of this registered replication report until the effects of verbal descriptions on the probative value of suspect IDs are more thoroughly understood. For the moment, it is not clear if psychologists should be advising the legal system to attach *more weight or less weight* to witnesses who identify a suspect after having provided a detailed verbal description of the perpetrator's face. Our further advice echoes a point made by Rotello, Heit, & Dubé (in press). In registered replications, the optimal approach may not be to insist that the original procedure be followed exactly, with no additional conditions included. With regard to the first registered replication, an argument could be made that a better way to have run the replication studies would have been to add a target-absent condition, with participants randomly assigned to each condition. Ignoring the false ID rates obtained from the target-absent condition provides an exact replication of the original study. But taking into account the additional information provided by the target-absent data would provide a better theoretical understanding of why verbal descriptions affect memory performance and how the field should advise the legal system.

References

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, S., Birch, S., Birt, A. R., ...Zwaan,

R. A. (2014). Registered replication report: Schooler & Engstler-Schooler

(1990). *Perspectives on Psychological Science, 9*, 556–578.

Clare, J., & Lewandowsky, S. (2004). Verbalizing facial memory: Criterion effects in

verbal overshadowing. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition, 30, 739-755. doi:10.1037/0278-7393.30.4.739

Dehon, H., Vanootighem, V., & Bredart, S. (2013). Verbal overshadowing of face

memory does occur in children too! *Frontiers in Psychology, 4,* 1-12.

10.3389/fpsyg.2013.00970

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics.

Oxford, England: Krieger.

Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide (2nd

ed.). Mahwah, NJ: Erlbaum.

Meissner, C. A. (2002). Applied aspects of the instructional bias effect in verbal

overshadowing. *Applied Cognitive Psychology, 16,* 911-928. doi:

10.1002/acp.918

Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal

overshadowing effect in face identification. *Applied Cognitive Psychology, 15,*

603-616. doi:10.1002/acp.728

Memon, A., & Rose, R. (2002). Identification abilities of children: Does a verbal

description hurt face recognition? *Psychology, Crime & Law, 8*, 229-242, doi:

10.1080/10683160208401817

Mickes, L., Moreland, M. B., Clark, S. E., & Wixted, J. T. (2014). Missing the

information needed to perform ROC analysis? Then compute *d'*, not the

diagnosticity ratio. *Journal of Applied Research in Memory and Cognition, 3,* 58-62*.*

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver Operating Characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous vs. sequential lineups. *Journal of Experimental Psychology: Applied, 18,* 361-376. doi:10.1037/a0030609

Rotello, C. M., Heit, E., & Dubé, C. (in press). When more data steer us wrong: Replications with the wrong dependent measure perpetuate erroneous conclusions. *Psychonomic Bulletin & Review*.

Sauerland, M., Holub, F. E., & Sporer, S. L. (2008). Person descriptions and person identifications: Verbal overshadowing or recognition criterion shift? *European Journal of Cognitive Psychology, 20,* 497-528. DOI: 10.1080/09541440701728417

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22,* 36-71. doi:10.1016/0010-0285(90)90003-M

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17,* 99-139.

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.

Wixted, J. T. & Mickes, L. (2014). A signal-detection-based diagnostic feature-detection model of eyewitness identification. *Psychological Review, 121,* 262-276.

Footnotes

[1]It is not clear which procedure is more relevant to the real world. Real eyewitnesses do not usually provide a verbal description of the perpetrator's face seconds before viewing a lineup, but they also do not usually provide a verbal description of the perpetrator's face seconds after witnessing a crime.

[2]Some indication of whether a criterion shift occurred in studies that only used target-present lineups can be obtained by examining the filler ID rates (i.e., the proportion of lineups in which a filler was identified in the two conditions). If filler ID rates actually increased in the verbal condition, then the reduced suspect ID rate in that condition would probably not be attributable to more conservative responding. However, in RRR1, filler ID rates also decreased significantly in the verbal condition. In RRR2, filler ID rates did not differ across conditions, a null result that might indicate the absence of a criterion shift. Note that filler ID rates were not analyzed in the replication report, but the relevant data were reported (allowing us to analyze them).

[3]Prior research has often found that, compared to simultaneous lineups, sequential lineups reduce the correct ID rate. Because sequential lineups induce more conservative responding, they reduce the false ID rate as well, often so much so that the probative value of an ID made from a sequential lineup exceeds that of an ID made from a simultaneous lineup (Steblay et al., 2011). Despite the increase in probative value, recent ROC analyses suggest that, in addition to inducing conservative responding, sequential lineups also reduce discriminability (Mickes et al., 2012).

Table 1. Data from Clare & Lewandowsky (2004) Experiment 1.

	Chooses Suspect	Chooses Filler	Does Not Choose
	Correct ID Rate	*Filler ID Rate*	*Miss Rate*
Target-present (Suspect Guilty)	Control: .80 Verbal: .63	Control: .13 Verbal: .09	Control: .07 Verbal: .28
	*False ID Rate**	*Filler ID Rate*	*Correct Rejection Rate*
Target-absent (Suspect Not Guilty)	Control: .13 Verbal: .08	Control: .77 Verbal: .48	Control: .23 Verbal: .52

*Because there was no designated innocent suspect, the false ID rate is estimated by dividing the filler ID rate by the lineup size (6) in accordance with standard practice. The correct and false ID rates for the verbal description condition were combined across their holistic vs. featural manipulation.