Effects of Verbalization on Lineup Face Recognition in an Interpolated Inspection Paradigm

KEITH B. LYLE* and MARCIA K. JOHNSON

Yale University, USA

SUMMARY

We examined the effects of describing a target face on face memory in a procedure in which subjects either did or did not inspect a single distractor face to determine if it matched the target face before attempting lineup identification of both the target and the distractor. Verbalization did not affect accuracy of identification of the target face; however, it increased accuracy during inspection of the distractor face and improved subsequent lineup identification of the distractor face. The absence of description-related impairment of memory for the target face (i.e., the failure to find verbal overshadowing), and the fact that verbalization has the potential positive effect of reducing false alarms to the distractor, points to the complex relation between verbalization and memory for faces.

Copyright © 2004 John Wiley & Sons, Ltd.

In the course of a criminal investigation, a single eyewitness may be called upon to verbally describe a perpetrator’s face, inspect one or more individual photos to determine whether they show the face of the perpetrator, and attempt photo lineup identification of the perpetrator’s face. Hereafter we will refer to these activities as description, inspection, and lineup identification. Because a single witness may be given multiple tests of face memory, it is important to ask whether participation in one test affects accuracy on another test. Therefore, various combinations of face memory tests have been investigated, including varying the presence of description prior to either lineup identification (e.g., Dodson, Johnson, & Schooler, 1997; Ryan & Schooler, 1998; Schooler & Engstler-Schooler, 1990) or inspection (Mauldin & Laughery, 1981; McClure & Shaw, 2002; Memon & Bartlett, 2002), and varying the presence of interpolated inspection after description and before lineup identification (Cutler, Penrod, & Martens, 1987; Dysart, Lindsay, Hammond, & Dupuis, Experiment 2, 2001; Lindsay, Nosworthy, Martin, & Martynuck, Experiment 1, 1994). These studies have shown that participation in one test can affect performance on another, as in the verbal overshadowing effect, whereby description relative to no-description impairs lineup identification (Schooler & Engstler-Schooler, 1990; Meissner & Brigham, 2001).

*Correspondence to: K. B. Lyle, Department of Psychology, Yale University, New Haven, CT 06520-8205, USA.
E-mail: keith.lyle@yale.edu

Contract/grant sponsor: National Institute on Aging; contract/grant number: AG09253.
Contract/grant sponsor: National Science Foundation.

Copyright © 2004 John Wiley & Sons, Ltd.
However, little is known about the effect of description on lineup identification in the presence of interpolated inspection. The question of how description and inspection interact to affect lineup identification was suggested by Schooler, Ryan, and Reder (1996), who, based on the results of a study by Ryan and Schooler (1994), proposed that description may moderate the potentially detrimental effect of viewing interpolated faces on lineup identification. In that study, subjects simply viewed four faces sequentially with the intention to remember them. Later, subjects attempted to identify each face from separate lineups. Verbal subjects described each face immediately after viewing it. Control subjects did not describe any of the faces. Thus, description varied while exposure to interpolated faces was constant. Overall, verbal subjects were more accurate than controls on all four lineups. Schooler et al. (1996) speculated that lineup identification was affected by both retroactive and proactive interference caused by viewing multiple faces. The authors suggested that description might have reduced the interference, resulting in higher accuracy. To the extent that inspecting faces is a special case of viewing multiple faces, inspected faces may interfere with a target face and vice versa, and description may counteract the interference.

Likewise, it is unclear how description might affect inspection itself. There are at least three published studies in which the presence or absence of some sort of description task was systematically varied prior to subjects inspecting individually presented faces (Mauldin & Laughery, 1981; McClure & Shaw, 2002; Memon & Bartlett, 2002). In those studies, description and no-description subjects did not differ significantly on inspection accuracy.

In Experiment 1 of this paper, we fully crossed a target-face description task with an interpolated-face inspection task in order to answer the following questions about the interplay of description, inspection, and lineup identification. One, does either description or inspection impair lineup identification of a target face? The type of inspection task we used is similar to a show-up procedure in which the witness is presented with the face of a single suspect and the witness must decide whether the suspect is the perpetrator. Two, do description and inspection interact to determine target lineup identification? Three, does description affect inspection, either by affecting whether an individually presented face is accurately classified as not matching a target face or by affecting incidental memory for an inspected face?

**EXPERIMENT 1**

Experiment 1 had a 2 (description or no description) × 2 (inspection or no inspection) between-subjects design.

**Participants**

Eighty Yale University students (52 females) participated either in partial fulfillment of a course requirement or for money.¹ Twenty subjects were randomly assigned to each of four conditions.

¹Forty six of the participants described themselves as Caucasian and 34 as non-Caucasian. Comparable numbers were obtained in Experiment 2. Preliminary analyses indicated that race did not affect performance on any of our measures in either experiment so we ignored race as a variable in all reported analyses.
Materials

We used the same photographs of faces as in Dodson et al. (1997, Experiment 2). Two of the photographs were of Caucasian males with dark hair. Each photo showed a frontal view of the face of one of the two men. There was a separate, corresponding lineup for each face that contained a novel photo of the face, which was also a frontal view but which was taken in a different setting, and five distractor faces. The distractors were of the same sex, race, and approximate age as the target and had similar hair colour and style. None of the individuals in either lineup appeared in the other lineup. The photos were projected from slides onto a blank white wall. Each of the two faces served as the Study and Comparison Face (see below) an equal number of times.

We also gave subjects a questionnaire developed by Dunning and Stern (1994), which asked, ‘How would you best describe your decision process when you chose a face from the lineup?’ The question was followed by eight statements. Subjects endorsed as many statements as applied. Following upon a prior report by Memon and Bartlett (2002) that self-report of basing lineup decisions on individual facial features (i.e. featural processing) was positively correlated with lineup identification accuracy, we focus on responses to the statement, ‘I based the judgment on specific facial features (e.g., nose, hair, eyes).’ Subjects in the inspection conditions filled out separate questionnaires in reference to the Study and Comparison Face lineups.

Procedure

All subjects began the experimental procedure by viewing the Study Face for 10 s. They were told to pay attention to the face and try to remember it for later in the experimental session. Next, as in previous verbal overshadowing studies (Dodson et al., 1997; Schooler & Engstler-Schooler, 1990), all subjects performed reading comprehension problems (Kaplan Educational Centers, 1998) as a distractor task.

After working on the reading comprehension problems for 5 min, subjects in the no-description conditions worked on additional reading comprehension problems for 5 more min, while subjects in the description conditions spent the next 5 min writing a description of the Study Face. Description subjects were given the following instructions:

Now we would like you to write a description of the person that you just saw. Please write a detailed description of the person’s face. Write the description so that someone reading the description would be able to identify the person. You will have five minutes to write the description. Try to use the full five minutes, providing as much detail as you can about the person’s facial features.

As soon as the allotted time was up for the reading comprehension or description task, subjects in the inspection conditions were told that they would see a second slide and that they should judge whether the individual in the second photo was the same person as in the first photo by marking either yes-same or no-different on a response sheet. The Comparison Face was then displayed until subjects responded. Since the Study and Comparison Faces were always different, the correct response was always no-different. In all, the inspection task took 1 to 2 min to administer.

2In both Experiment 1 and Experiment 2, we tested the effects of our manipulations only on lineups that included a photo of the to-be-identified individual, instead of also using target-absent lineups, because we were interested primarily in whether inspection produced a form of interference that might reduce witnesses’ ability to recognize a previously seen individual.
For subjects in the no-inspection conditions, the reading comprehension or description task was followed by a 2-min break, during which subjects completed some lab paperwork. Subjects in the inspection conditions took this break immediately after completing the inspection task. After the 2-min break, subjects in the inspection conditions were informed that the Study and Comparison Faces were, in fact, the faces of two distinct individuals.

Next, all subjects were given instructions for the lineup task. Subjects in the no-inspection conditions were told that they would see a set of six photographs of faces and that they should look for the person they had seen in the earlier slide. The lineup identification procedure for inspection subjects was slightly different in that they saw two separate lineups, one each for the Study and Comparison Faces. Half of the inspection subjects were given the Study Face lineup first and the Comparison Face lineup second. The order was reversed for the other half of the inspection subjects. Subjects were told which face they should look for in each lineup (i.e. the first or the second face that they saw). Subjects in all conditions were told that a photo of the person they were looking for was in the lineup, so they should make a choice even if they were guessing. Responding that the individual was not present was not an option. Lineups were displayed until subjects selected a photo.

After completing the lineup task, subjects were given a decision process questionnaire for each lineup that they had seen. Subjects in the inspection conditions first were given a questionnaire for the first lineup presented and then were given a second, identical questionnaire for the second lineup presented.

Results and discussion

For statistical comparisons, alpha was set at the conventional level of 0.05, except where otherwise noted.

Table 1 shows the percentages of correct responses on the lineup tests as a function of whether subjects engaged in description, inspection, both, or neither.

Our first question was whether either description or inspection impaired Study Face lineup identification. A $2 \times 2 \times 2$ (description or no description) $\times$ (inspection or no inspection) $\times$ (correct or incorrect response) log-linear analysis on the Study Face lineup data indicated that there was no main effect of description; an equal number (37.5%) of description and no-description subjects identified the Study Face. Fewer inspection (30%) than no-inspection (45%) subjects correctly identified the Study Face, but the main effect

Table 1. Percentages of correct responses on the inspection task, Study Face lineup, and Comparison Face lineup in Experiment 1 depending on whether subjects described the Study Face and whether they inspected the Comparison Face (frequencies in parentheses)

<table>
<thead>
<tr>
<th>Description?</th>
<th>Inspection?</th>
<th>Inspection task</th>
<th>Study Face lineup</th>
<th>Comparison Face lineup</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
<td>—</td>
<td>40 (8)</td>
<td>—</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>75 (15)</td>
<td>35 (7)</td>
<td>20 (4)</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>—</td>
<td>50 (10)</td>
<td>—</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>100 (20)</td>
<td>25 (5)</td>
<td>50 (10)</td>
</tr>
</tbody>
</table>

$^3$Order of lineup presentation did not interact significantly with the independent variables of primary interest in either Experiment 1 or Experiment 2, therefore we ignore it in our presentation of the results.

Copyright © 2004 John Wiley & Sons, Ltd.

of inspection also was not significant ($G^2(1) = 1.92, p > 0.05$). Our second question was whether description interacted with inspection to affect Study Face lineup identification. The Study Face lineup data did not reveal a significant three-way interaction between description, inspection, and accuracy ($G^2(4) = 2.8, p > 0.05$).

Thus, analysis of the Study Face data failed to reveal a verbal overshadowing effect using young adult subjects and standard description instructions. That is, we did not find poorer performance on the Study Face lineup in the description-only group (50%) than in the no-description/no-inspection group (40%). This is perhaps not surprising given that the effect size is small according to a recent meta-analysis (Meissner & Brigham, 2001) and there have been other failures to replicate (S. B. Lovett, M. Y. Small, & S. A. Engstrom, poster presented at the annual meeting of the Psychonomic Society, St. Louis, MO, 1992; Meissner, Brigham, & Kelley, 2001; Memon & Bartlett, 2002).

Nor did we find that inspection of the Comparison Face reduced identification of the Study Face. This bolsters the conclusion of Dysart et al. (2001), who previously argued that inspection does not interfere with target face lineup identification, although, in their study, unlike in ours, all subjects described the target face, leaving open the possibility that description counteracted interference that they otherwise would have observed (see Schoon et al., 1996).

Our next question was whether description affected accuracy in our inspection task. Column 3 of Table 1 shows that more description (100%) than no-description (75%) subjects correctly judged that the Comparison Face was not the same as the Study Face ($X^2(1) = 3.66, p < 0.06$). The positive effect of description on the inspection task is consistent with previous studies in which description non-significantly improved recognition responses to individually presented faces (Mauldin & Laughery, 1981; McClure & Shaw, 2002).

We then asked whether description of the Study Face affected incidental memory for the Comparison Face as measured by a lineup test. We found more correct Comparison Face lineup identifications among subjects who described the Study Face (50%) than among those who did not (20%). This difference approached the conventional level of significance ($X^2(1) = 2.75, p < 0.10$).

Finally, to assess whether description affected the relationship between Study Face lineup identification accuracy and self-report of basing identification on individual facial features, we correlated whether or not subjects endorsed the statement ‘I based the judgment on specific facial features (e.g., nose, hair, eyes)’ on the Study Face lineup decision questionnaire with whether or not they correctly identified the Study Face. For no-description subjects, there was a positive correlation between self-report of featural processing and lineup identification accuracy ($r(40) = 0.39, p < 0.05$). Description subjects showed a markedly different pattern in which the two measures were non-significantly negatively correlated ($r(40) = -0.27, p < 0.09$).4

Memon and Bartlett (2002) previously reported that feature use was positively related to accurate lineup identification. However, because of a relatively small $n$, Memon and Bartlett collapsed across description and no description conditions. Our results qualify theirs in that the positive relationship between feature use and accuracy did not hold when the target face was described. Why might description eliminate the relationship between feature use and accuracy? One possibility is that description subjects base their lineup decisions on what they remember of their description of the face. To the extent that

---

4The relationship between feature use and accuracy was not affected by inspection for either description or no-description subjects.
description subjects may generate non-veridical information about facial features, an idea originally proposed by Schooler and Engstler-Schooler (1990) and recently elaborated on by Meissner et al. (2001), relying on a description to make a lineup identification could lead to inaccurate choices. On the other hand, given that at least some elements of the description may be veridical, using descriptions could be helpful. Moreover, some description subjects may rely on their visual memory of the facial features, even though they generated a verbal description of them. Variability induced by the fact that not all verbal representations are misleading or even used may make it difficult to see a clear relationship, positive or negative, between feature use and lineup identification accuracy in description subjects.

We obtained some evidence for two novel effects of describing a target face on subsequent inspection of a non-target face. There were trends toward description increasing correct rejections of the Comparison Face \((p < 0.06)\) and increasing correct lineup identifications of the Comparison Face \((p < 0.10)\). However, because neither of these effects reached conventional levels of significance, we withhold discussion of them until after demonstrating that the effects were replicated in a second experiment.

**EXPERIMENT 2**

In Experiment 2, we show that the beneficial effects of description on the inspection task and the Comparison Face lineup were replicated and generalized to a different set of materials. In addition, we examine a new variable: retention interval between inspection and lineup tests. In Experiment 2, all subjects performed the inspection task but half of the subjects were given the lineups 2 min after inspection and half were given the lineups 2 days after inspection. The manipulation of retention interval is of interest for two reasons. First, in real-world settings, while the same witness may engage in both inspection and lineup identification, the two tests need not be given in temporal proximity. Second, the detrimental effect of description on lineup identification (i.e. verbal overshadowing) has been found to endure as long as 2 days (Schooler & Engstler-Schooler, 1990) or as briefly as 24 min (Finger & Pezdek, 1999). We ask whether the beneficial effect of description on Comparison Face lineup identification is relatively temporary or long lasting.

Experiment 2 had a 2 (description or no description) \(\times 2\) (short or long retention interval) between-subjects design.

**Participants**

Eighty Yale University undergraduates (52 females) participated in partial fulfilment of a course requirement. Twenty subjects were randomly assigned to each of four conditions.

**Materials**

As in Experiment 1, we needed two faces and corresponding lineups. Like other researchers (Collishaw & Hole, 2000; Perfect, Hollins, & Hunt, 2000), we obtained photographs of faces from the University of Stirling’s picture database (http://pics.psych.stir.ac.uk). The Study and Comparison Faces were Caucasian males with light hair. The pictures showed three-quarter profile views of the faces. For each face we created a corresponding lineup that contained a novel view of the face and five distractor faces, all
seen frontally. The distractors were similar to the targets in terms of sex, race, age, hair colour, and hairstyle. No individual appeared in both lineups. All photographs were displayed on a computer monitor.

**Procedure**

Presentation of the Study Face, the reading comprehension distractor task, and the description task were identical in all conditions to the procedure followed for description and control subjects in Experiment 1, save that exposure time to the Study Face was 5 s instead of 10.

After time expired for description subjects to perform the description task and for control subjects to perform the distractor task, all subjects were given the inspection task. The inspection task was the same as in Experiment 1 except that it was performed on a computer, which recorded the total time that each subject was exposed to the Comparison Face.

Following inspection of the Comparison Face, there were two retention intervals. Half of the subjects were randomly assigned to the short delay condition in which the lineup tests were given 2 min after the inspection task. The other half of the subjects were in the long delay condition, in which they were given the lineups at a second session, 48 h after the inspection task. As in Experiment 1, subjects were informed, immediately before the lineups were administered, that the Study and Comparison Faces were the faces of two different individuals.

Administration of the Study and Comparison Face lineups was the same as in Experiment 1, with one exception. Although the Study and Comparison Faces were always present in their respective lineups, subjects in Experiment 2 were given a Not Present option so they could indicate that they did not think that the person they were looking for was in the lineup.

**Results and discussion**

Four subjects, one in each condition, were excluded from all analyses because they misunderstood the inspection task instructions.

Table 2 shows the percentages of correct responses on the inspection and lineup tests as a function of whether subjects engaged in description and when they were given the lineups.

Based on the results of Experiment 1, we predicted that more subjects in Experiment 2 would respond accurately on the inspection task when they had described the Study Face than when they had not. Since the manipulation of retention interval occurred after the
inspection task, we were able to combine the data from subjects in the short and long conditions in the analysis of inspection accuracy. The third column of Table 2 shows that, as predicted, a larger percentage of description (95%) than control (79%) subjects correctly rejected the Comparison Face as differing from the Study Face ($X^2(1) = 2.88, p < 0.05$, one-tailed).

Our next prediction based on the results of Experiment 1 was that description subjects at the short retention interval would exhibit greater accuracy on the Comparison Face lineup than would no-description subjects at the same retention interval. We performed a 2 (description or no description) $\times$ 2 (short or long retention interval) $\times$ 2 (correct or incorrect response) log-linear analysis on the Comparison Face lineup data. There was a main effect of description ($G^2(1) = 4.38, p < 0.05$) such that description subjects (69%) were more accurate than control subjects (45%), but the main effect was qualified by a significant three-way interaction ($G^2(4) = 9.88, p < 0.05$). The interaction arose because a larger percentage of subjects in the description-short condition (84%) than in the control-short condition (37%) correctly identified the Comparison Face ($X^2(1) = 7.05, p < 0.01$, one-tailed), but an equal percentage (53%) of description-long and control-long subjects identified the Comparison Face. Thus, description appeared to improve Comparison Face lineup identification after 2 min but not after 48 h.

In Experiment 2, we recorded exposure time to the Comparison Face for each subject. Description (15.0 s) and no-description (14.3 s) subjects did not differ in the average amount of time they took to make their decision in the inspection task ($t(73) = 0.40, p > 0.05$). Therefore, description subjects’ superior performance on the inspection task and Comparison Face lineup was not due to those subjects being exposed to the Comparison Face for longer than were control subjects.

GENERAL DISCUSSION

These studies explored ways in which participating in tests of face memory might affect performance on other tests. We found that neither target-face description nor inspection of an interpolated face, like a suspect’s face during a show-up, interfered with later lineup identification of a target face.

Although description had no effect on target-face lineup identification accuracy, description increased the accuracy of inspection. Subjects were more likely to correctly reject an individually presented distractor face as being the same as the target face following description versus no-description. It is well established that presenting lineup faces sequentially for individual inspection produces fewer false positive identifications than does simultaneously displaying all of the lineup members (e.g. Lindsay et al., 1991; Lindsay & Wells, 1985; Sporer, 1993; Wells, 1984; Wells et al., 1998). Our finding is important to follow up on because it suggests that describing a target face prior to inspecting singly presented faces may reduce false alarms even further.

We also found that description improved lineup identification of the distractor face that subjects inspected, but only when memory was tested 2 min after inspection, rather than 2 days. Schooler et al. (1996) suggested that faces may proactively interfere with lineup identification of subsequently viewed faces and that description may reduce the interference. Although, in an inspection paradigm, a no-interference control condition cannot be included, because subjects cannot inspect a face without previously having viewed a potentially interfering target face, the pattern of results we obtained suggests that memory
for the Comparison Face suffered from some sort of temporary proactive interference from the Study Face and that describing the Study Face reduced the interference. For example, the Study Face, which subjects were actively trying to remember, may have been more strongly associated with retrieval cues during the lineup tests than was the Comparison Face, at least at a short retention interval. According to a blocking account of interference in recognition memory (Chandler & Gargano, 1998; Windschitl, 1996), this might have caused the Study Face to be inadvertently retrieved during test of the Comparison Face, thereby blocking retrieval of the Comparison Face. Description may reduce blocking by increasing the difference between the memorial representations of the Study and Comparison Face thereby increasing the likelihood that any particular cue is specific to one of the faces (Johnson, Hashtroudi, & Lindsay, 1993).

In two previous studies (Brown & Lloyd-Jones, 2003; Dodson et al., 1997), description of one face was found to impair recognition memory for faces other than the one that was described. This makes it particularly surprising that we found that describing one face improves identification of another face. However, in the earlier studies, the non-described face was encountered before description, while our subjects encountered it after description. This methodological difference suggests an alternative explanation for description subjects’ higher Comparison Face lineup accuracy based on the idea of transfer-inappropriate processing (Schooler, Fiore, & Brandimonte, 1997). Verbalization may affect face recognition by creating a mismatch between the way faces are encoded versus retrieved. Faces are assumed to be encoded and retrieved relatively holistically but verbalization induces subjects to attempt to identify faces by retrieving feature information (i.e. the type of information favoured in verbal descriptions). This suggests that, following description, test faces would be incidentally encoded by feature, because their features would be compared to the features of the Study Face, while, in the absence of description, incidental encoding would more likely be holistic.

Therefore, if memory for a test face were itself tested, as it was in our procedure (i.e. on the Comparison Face lineup), there would be a high probability that both encoding and retrieval would be feature based for description subjects. For control subjects, there might be more variability between encoding and retrieval because, for example, although retrieval may be holistic for most subjects, some may attempt feature-based retrieval. This would lead to a larger number of description than no-description subjects correctly identifying the test face, at least at a short retention interval, when feature information about the test face presumably is most likely to be available for retrieval.

This alternative account highlights the potential importance of controlling for the way in which faces are encoded and retrieved in future studies of verbalization and face recognition. Our correlational data from Experiment 1, which showed that verbalization affected the relationship between feature-based retrieval and lineup identification accuracy suggests that, in order to fully understand the complex relationship between verbalization and face recognition, it may be useful to exert more control over the nature of processing.

ACKNOWLEDGEMENTS

This research was supported by National Institute on Aging Grant AG09253 to MKJ and a National Science Foundation Graduate Fellowship to KBL. We thank Elizabeth Allen and Joseph Bates for their help in conducting Experiment 1. We also thank Christian Meissner and Jonathan Schooler for their helpful comments on earlier versions of this manuscript.
REFERENCES


