Journal of Experimental Psychology: Applied

Do Masked-Face Lineups Facilitate Eyewitness Identification of a Masked Individual?

Krista D. Manley, Jason C. K. Chan, and Gary L. Wells

Online First Publication, December 17, 2018. http://dx.doi.org/10.1037/xap0000195

American Psychological Association

Journal of Experimental Psychology: Applied

1076-898X/18/$12.00 http://dx.doi.org/10.1037/xap0000195

Do Masked-Face Lineups Facilitate Eyewitness Identification of a Masked Individual?

Krista D. Manley, Jason C. K. Chan, and Gary L. Wells
Iowa State University

Perpetrators often wear disguises like ski masks to hinder subsequent identification by witnesses or law enforcement officials. In criminal cases involving a masked perpetrator, the decision of whether and how to administer a lineup often rests on the investigating officer. To date, no evidence-based recommendations are available for eyewitness identifications of a masked perpetrator. In 4 experiments, we examined lineup identification performance depending on variations in both encoding (studying a full face vs. a partial/masked face) and retrieval conditions (identifying a target from a full-face lineup vs. a partial/masked-face lineup). In addition, we manipulated whether the target was present or absent in the lineup in Experiments 3 and 4. Across all experiments, when participants had encoded a masked face, the masked-face lineup increased identification accuracy relative to the full-face lineup. These data provide preliminary evidence that matching lineup construction to how witnesses originally encoded the perpetrator may enhance the accuracy of eyewitness identifications.

Public Significance Statement

This study provides preliminary evidence that matching the facial appearance of lineup members to how witnesses originally encoded the perpetrator may enhance eyewitness accuracy. Additionally, it shows that witnesses are faster to respond and more confident in target identifications when the lineup matches encoding compared with when it does not.

Keywords: eyewitness identification, transfer-appropriate processing, disguise, lineup construction

Supplemental materials: http://dx.doi.org/10.1037/xap0000195.supp

Imagine as a detective, you learn a woman was assaulted and robbed outside of her residence by a man wearing a ski mask. Even though the ski mask only revealed his eyes, the victim was in close contact with the perpetrator. Shortly after the crime, a man was captured and identified as the suspect. As the lead investigator in the case, it is up to you to make the tough decision of how or whether to administer a lineup. Should the victim be shown a lineup, given that she only saw the perpetrator's eyes? If so, how should the lineup be constructed?

The above scenario provides an example of the difficult questions faced by the criminal justice system when a crime involves the eyewitness identification of a masked perpetrator—given the dearth of empirical data, there are few evidence-based policy recommendations regarding how lineups should be administered for these situations. For example, the 2012 New Jersey Supreme Court's landmark guidelines for jurors included only a simple warning that "the perpetrator's use of a disguise can affect a witness's ability both to remember and identify the perpetrator. Disguises like hats, sunglasses, or masks can reduce the accuracy of an identification" (Report of the Special Master, *State v. Henderson*, 2011; Innocence Project, 2012a, 2012b). Although these instructions inform jurors that concealment-based disguises may impair eyewitness identifications, they provide no procedural recommendations. In the example described in the opening paragraph, because the victim did not see the full face of the perpetrator, investigators may choose to withhold a photo lineup from the victim. However, because lineup administration decisions are left to the discretion of the investigator, a photo lineup may be administered—which has been the case in similar situations (*People v. Williams*, 1979; *State v. Courteau*, 1983; *State v. Fierro*, 1971). Given that investigators do sometimes administer a lineup when a masked perpetrator is involved, it is important to examine the circumstances under which witnesses are more or less effective at such lineup identifications. We note, however, eyewitness identifications of a masked perpetrator go beyond the confines of a lineup. Even if the investigating officer decides to withhold a lineup from a witness, the witness may eventually be asked to identify the perpetrator in court (e.g., a witness may be asked "Is the person who did this to you in the courtroom today?") when the suspect is no longer wearing a mask.

Krista D. Manley, Jason C. K. Chan, and Gary L. Wells, Department of Psychology, Iowa State University.

This research was partially supported by the Wendy and Mark Stavish fund to Gary L. Wells.

Correspondence concerning this article should be addressed to Krista D. Manley, Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, IA 50011. E-mail: kmanley@iastate.edu

In the present study, we begin by examining whether presenting a lineup of faces that match the perceptual appearance of the target (e.g., a lineup consisting of masked faces for a participant who initially saw a masked face), as opposed to a lineup of faces that do not match the perceptual appearance of the target (e.g., a lineup consisting of unmasked faces for a participant who initially saw a masked face), would lead to superior identification performance—even if the administration of such lineups requires one to partially conceal the faces in a lineup, which in turn reduces the perceptual features available at retrieval for the eyewitness. The theoretical framework of transfer-appropriate processing provides the rationale for this prediction. The transfer-appropriate processing framework states that performance on a memory test is best when the processes activated at retrieval match those at encoding (Morris, Bransford, & Franks, 1977).

Transfer-Appropriate Processing in Eyewitness Identification

The eyewitness identification literature is replete with examples that are broadly consistent with the general framework of transfer-appropriate processing. For example, reinstating the original emotional or environmental context during retrieval often increases correct identifications and reduces false identifications compared to when context differs between encoding and retrieval (e.g., Chandler & Fisher, 1996; Shapiro & Penrod, 1986; Smith & Vela, 2001). Of particular relevance to the present study, Davies and Flin (1984) investigated whether matching the appearance of lineup members to encoding can enhance eyewitness identification for distorted faces. Here, the perpetrator's face was either distorted by a tight-fitting, semitransparent nylon stocking mask or not. Participants were first shown four pictures of men for 30 s each wearing a stocking mask or no mask. Next, participants attempted to identify the four targets from either an array of 16 individuals shown wearing stocking masks or shown wearing no masks. Identification accuracy was better when the lineup matched what happened at encoding than when it did not. Using a different form of concealment, Hockley, Hemsworth, and Consoli (1999) investigated the impact of sunglasses on face recognition. In their Experiments 2 and 3, participants encoded either faces wearing sunglasses or not. They were then given a recognition test that consisted of both old and new faces with and without sunglasses. When participants studied faces that did not have sunglasses, they showed a higher hit rate and lower false alarm rate toward faces not wearing sunglasses compared to faces wearing sunglasses. However, when participants studied faces that were partially covered by sunglasses, they showed a higher hit rate (and a comparable false alarm rate) for faces with sunglasses than those without.

Although few studies have examined the impact of concealment on identification accuracy (e.g., Cutler, Penrod, & Martens, 1987b; Davies & Flin, 1984; Mansour et al., 2012), it is well-known that perpetrators often wear masks when committing a crime (Noble, 2013). Several mechanisms might explain why wearing a mask is effective at reducing identification accuracy. For instance, a mask can cover or obscure the distinctive features of a face (e.g., a scar, facial hair, a mole, etc.), decrease the number of facial cues available for encoding (Brewer, Weber, & Semmler, 2013), and/or prevent holistic processing of the face (Farah, Tanaka, & Drain, 1995). Here we consider results from the face recognition literature that may help to understand the mechanisms by which concealment disrupts identification accuracy.

Masking and Face Processing

According to the holistic processing account of face perception, faces are represented as an integrated whole rather than as separate featural components (i.e., featural processing; Tanaka & Farah, 1993; for reviews see Bruce & Young, 2012; Maurer, Le Grand, & Mondloch, 2002; Tanaka & Simonyi, 2016). Proponents of the holistic processing approach suggest that findings like the face-inversion effect (Yin, 1969; for review see Valentine, 1988) and the composite-face effect (Young, Hellawell, & Hay, 1987) are consistent with the idea that faces, unlike nonface objects (e.g., houses), are particularly sensitive to the configural relations among features rather than the individual physical attributes (e.g., size, shape) of the features—the hallmark of holistic processing. The face inversion effect (Yin, 1969) is the finding that people's ability to discriminate between similar faces is dramatically reduced when the faces are inverted, a deficit not found with nonface objects. It is thought that faces cannot be effectively processed holistically (a potential requirement of accurate face recognition) when inverted.

Another finding that supports the idea that faces are processed holistically while other objects (e.g., houses) can be processed as individual attributes comes from change-blindness experiments. In a simple *change detection* task, people perform better at deciding whether or not there was a change to a face (e.g., a different jaw) than to a house (e.g., different porch or door). However, when the task requires the person to correctly name *which* feature changed, people perform better with houses than faces (Wilford & Wells, 2010). These results are consistent with the idea that holistic processing is efficient at detecting configural changes but perhaps inefficient at localizing which local or featural changes—as a result, when the task requires one to identify a pair of eyes, presenting a full-face lineup, which presumably invites holistic processing, might be detrimental to identification.

In addition to the inversion effect and the change detection effect, there is a composite-face effect that tends to support the idea that faces are processed holistically. The composite-face effect (Young et al., 1987) is found using "composites" for which the top half of one face is combined with the bottom half of another face, thereby creating a perceptually new face. When the halves are misaligned, people are quicker to recognize the face as familiar, but when they are aligned recognition speed suffers. This is because people are unable to ignore the holistic representation when the halves are aligned, leading to decreased recognition accuracy and speed compared to when the halves are misaligned or when the composite is inverted (Tanaka & Simonyi, 2016).

Of particular relevance to the present study are two findings. First, Tanaka and Farah found that after participants have encoded a whole face, they were better at recognizing individual features of that face if those features were presented as part of a whole face than when those features were presented in isolation, an effect termed the whole-face advantage (also see Tanaka & Sengco, 1997). In contrast, Leder and Carbon (2005) found that when participants encoded isolated facial features, they were less likely to recognize those features when they were presented in the context of a whole face than when they appeared in isolation, a finding termed whole-face interference (similar to the composite-face effect; Young et al., 1987). Leder and Carbon (2005) argued that during recognition, face perception automatically engages holistic processing even when the existing memory representation consists of isolated features. This automatic holistic processing thus interferes with successful completion of the feature recognition task, which demands featural processing for an accurate match to memory.

Taken together, the extant findings suggest that matching the appearance of the lineup members to encoding may enhance identification of a masked face (which is similar to presenting a feature in isolation). Although this idea seems intuitive at first glance, a deeper examination reveals that the relation between transfer-appropriate processing and concealment-based masking is nuanced because concealment differs from other forms of disguise (e.g., distorting a face with a tight-fitting see-through nylon stocking) from the perspective of face processing in two important ways. First, the features in an undistorted face differ from those in a distorted face. For example, a see-through nylon stocking can alter the appearance of a face by stretching/flattening its features and compressing its contours. Therefore, it is perhaps not surprising that one would have trouble recognizing a distorted face when presented with its undistorted version. This alteration in appearance does not apply in the case of concealment-based masking, in which the visible features (e.g., eyes) of an unmasked face are identical to those of that same face when it is masked. Second, during lineup presentation, a masked face presents fewer features to participants than an unmasked one, so a masked-face lineup actually reduces the number of features available at retrieval than a full-face lineup—a difference that does not apply to distortion-based disguises.

How might matching the lineup appearance to encoding benefit identification performance? Here we invoke the idea that faces are typically processed holistically. Therefore, if a participant had encoded a full face (presumably holistically), presentation of a full-face lineup would constitute a lineup that is transfer-appropriate because participants would likely process the faces at both the encoding and the lineup holistically. In contrast, if a participant had encoded a masked face *with only the eyes visible*, which is presumed to involve featural instead of holistic processing (Farah et al., 1995), then presenting a lineup of full faces (which would be processed holistically) would be considered transfer-inappropriate, potentially leading to whole-face interference (Leder & Carbon, 2005).

Overall Design

All of the present experiments used a 2 (Encoding: full face vs. partial/masked face) × 2 (Retrieval: full-face lineup vs. partial-/masked-face lineup) within-subjects design, with participants completing one trial per condition (i.e., encode full face with a full-face lineup, encode full face with a partial-face lineup, encode partial face with a full-face lineup, encode partial-face with a partial-face lineup) for a total of four trials. The order of the trials was counterbalanced across participants by permutation. Each trial contained four phases. First, participants saw a target stimulus (i.e., partial face/masked face or full face). Second, they watched a 5-min *preidentification video*. Third, they attempted to identify the target from a three-face lineup (i.e., partial-face/masked-face lineup or full-face lineup). Fourth, participants watched a 5-min, *postidentification video*. In Experiment 1, we manipulated whether participants saw whole or partial faces at both encoding and retrieval. Experiment 2 was similar to Experiment 1 except that participants encountered masked faces instead of partial faces. In Experiment 3, we examined performance for target-absent as well as target-present lineups. In Experiment 4, we held all noncritical facial features in the full-face lineups constant, so that all faces were identical with the exception of the eyes. We delay descriptions of logic of Experiment 4 until its introduction.

Experiment 1

Method

Participants. Participants were 196 undergraduate students from Iowa State University. One participant did not follow instructions and was excluded from all analyses; this left 195 participants in the final data set. Given the novelty of our investigation, we did not have an established literature from which to estimate effect sizes. As a result, we adopted a rough estimate of required sample size based on previous experience that approximately 100 participants would be needed to detect a small effect in a single trial eyewitness identification experiment. To be on the conservative side in detecting the requisite Encoding × Retrieval interaction, we opted to double the sample size to approximately 200 participants (Simonsohn, 2014). Lastly, because we assigned our face materials across conditions based on a permutation design, we needed multiples of 24 participants to complete each counterbalance. As a result, we aimed to test approximately 192 participants in Experiment 1. The Iowa State Institutional Review Board approved the research protocols (IRB #15–663).

Materials and procedure. Twelve faces were generated using the Faces™ 4.0 application (IQ Biometrix, Inc., 2003). All of the faces had different eyes, eyebrows, noses, and lips, with the exterior features held constant. We selected four categories of eye shapes based on the labels in the Faces™ 4.0 software: narrow, heavy, almond blue, and bulging. Three pairs of eyes were selected from each category. We then paired stimuli from each eye category with noses, lips, and eyebrows from a consistent category (e.g., the faces with narrow eyes all had a nose from the "average pointed" category; see Figure 1). This was done to ensure that all faces in a lineup would look similar. The stimuli in each eyes' category were randomly assigned to the four encoding-retrieval conditions. The partial-face stimuli were created by cropping the full-face stimuli to only the area around the eyes (see Figure 1). Stimuli were counterbalanced such that each served as the target across conditions equally often. The same target picture appeared at study and test, either cropped to show just the eyes or not cropped to show the whole face.

Eight videos approximately 5 min long (M = 4:53 min; range = 4:31–5:21 min) each were used as preidentification and postidentification videos. The four preidentification videos were dialogue-free, had music, and depicted short stories (e.g., a theft, kindness, silent communication, and money transfer). The four postidentification videos contained dialogue and each described a basic science topic (e.g., neurological effects of caffeine, biology of sharks, biology of dogs, and the human digestive system). Assignment of the videos to the trials was determined randomly for each participant. Each trial contained the four phases as described previously.

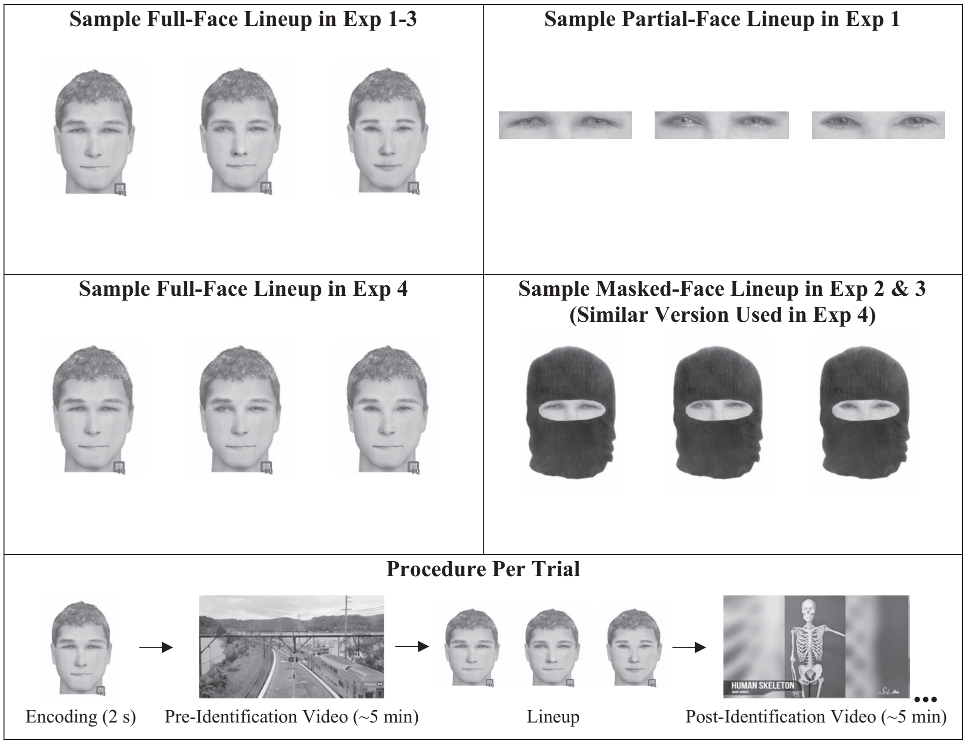

Figure 1. Examples of the stimuli used in Experiments 1–4 (not shown to scale). Note in Experiments 3 and 4 the "not present" option appeared centered under each lineup. In addition, the masked-face lineup for Experiment 4 would look similar to the masked-face lineup for Experiments 2 and 3 using the Experiment 4 stimuli. Used with permission by IQBiometrix Inc.

At the beginning of the experiment, participants read the following instructions for a cover story: "We are interested in how people remember different types of visual information (e.g., pictures of faces, videos with and without dialogue). For the pictures, please try your best to remember the eyes of the people we show you. For the videos, try your best to remember the content." During each trial of the experiment, participants first read the instructions: "You will now see a picture. Try to remember the eyes." Next, a partial face or a full face was displayed for 2 s. This presentation rate was determined following pilot testing to ensure nonceiling identification performance. Participants then watched a ~5 min preidentification video and completed a two-question multiple-choice test for the video (e.g., How did the main characters communicate? What was the woman listening to in the earphones?). Afterward, participants received a three-person target-present lineup. The full faces or the partial faces were distributed horizontally with numbers printed below each face. Only three faces were included in the lineup because we had a limited ability to generate unique faces that were not too similar.

Above each lineup, participants were given the following instructions: "We will now test your memory for the picture you saw earlier. Please select the number that corresponds to the correct answer." The target position was determined randomly. After participants made their recognition decision, they then rated their confidence between 0 and 100, with 0 = *not confident at all (i.e., a pure guess)* and 100 = *absolutely positive*. Following the lineup, participants watched a ~5 min postidentification video and then completed a two-question multiple-choice test for the video. The postidentification video was designed to reduce between-trial interference of the facial stimuli. After this, the next encoding trial was presented.

Participants were not told that they would complete four trials, that they would see some full faces and some partial faces, or that they would receive either full-face or partial-face lineups. Instead, under the premise that we were interested in how people remember different types of visual information, they were simply told to remember the eyes when they saw a picture, and to remember the content when they saw a video. This cover story allowed us to hold the encoding and lineup instructions constant across all trials for two purposes. First, we did not want participants to change their encoding strategy across trials, so we devised encoding instructions that did not divulge whether participants would see a full face or a partial face on that trial. Second, we did not want participants to know whether they would encounter a full-face or partial-face lineup on any given trial, which might also alter their encoding strategy. The entire task took approximately 60 min, was self-administered, and paced by the computer program.

Results and Discussion

The alpha level for all statistical analyses was set at .05 and all reported CIs are 95% CIs. Because each within-subjects condition contained a single trial of a binary outcome (correct vs. incorrect), we used generalized estimating equations (GEE) to analyze the effects of encoding and retrieval conditions on identification accuracy. For the dependent measures of confidence and response latency, linear mixed models (LMM) were used in lieu of an analysis of variance (ANOVA) because the latter would exclude participants who did not have a data point (e.g., confidence) for all conditions. For example, when examining confidence for correct identifications, a repeated measures ANOVA would only include participants who correctly identified the target across all four trials, thus severely reducing power. Estimates of discrimination (d') were derived using Luce's (1963) model (see Table 1). Note that response bias cannot be computed for Experiments 1 and 2 because participants were forced to choose during the lineup identification task. In all Results sections, we first report identification accuracy; we then report the data on confidence and response latency. In addition, materials for each experiment including stimuli and data can be found on the Open Science Framework (OSF; https://osf.io/5h4sg/?view_only=42807b9cb45844edb5458f25168ae45a). For each experiment, confidence ratings and response latency data for foil IDs and incorrect rejections can be found in the online supplemental materials.

Target identifications. Target identification was analyzed with a repeated-measures, 2 (Encoding: partial face vs. full face) × 2 (Retrieval: partial-face lineup vs. full-face lineup) GEE. Most importantly, there was a crossover interaction between encoding and retrieval conditions, $\chi^2(1, 194) = 10.84$, $p = .001$, $OR = 3.16$, CI [1.59, 6.27]. Specifically, when participants had encoded a full face, the full-face lineup produced more target identifications ($M = .81$, $SE = .03$, $d' = 1.52$) than did the partial-face lineup ($M = .71$, $SE = .03$, $d' = 1.12$). However, when participants had encoded a partial face, the full-face lineup produced fewer target identifications ($M = .68$, $SE = .03$, $d' = 1.02$) than did the partial-face lineup ($M = .79$, $SE = .03$, $d' = 1.43$). Therefore, consistent with the prediction based on transfer-appropriate processing, matching encoding and retrieval conditions improved eyewitness identification performance. Neither encoding nor retrieval conditions produced statistically significant main effects (see Table 1). Specifically, similar identification performance was observed regardless of whether participants encoded partial faces ($M = .73$, $SE = .02$) or full faces ($M = .76$, $SE = .02$), $\chi^2(1, 194) = .65$, $p = .42$. In addition, identification performance did not differ regardless of whether participants retrieved from a partial-face lineup ($M = .75$, $SE = .02$) or a full-face lineup ($M = .74$, $SE = .02$), $\chi^2(1, 194) = .003$, $p = .96$.

Confidence. The effects of encoding and retrieval conditions on confidence associated with correct and foil identifications were analyzed with separate LMMs. For *target identifications*, there was a crossover interaction between encoding and retrieval conditions, $F(1, 578) = 6.65$, $p = .01$ (see Table 1). When participants encoded a partial face, they were more confident in their target identifications made from a partial-face lineup ($M = 81.18\%$, $SE = 1.57$) than from a full-face lineup (M = 77.73%, $SE = 1.95$). However, when participants encoded a full face, they were less confident in their target identifications made from a partial-face lineup ($M = 74.26\%$, $SE = 1.97$) than from a full-face lineup ($M = 79.91\%$, $SE = 1.59$). We also examined the influence of encoding and retrieval on the confidence ratings associated with *foil identifications*. Here, neither the main effects nor the interaction between encoding and retrieval were significant, $F\text{s} < 1.22$, $p\text{s} > .27$.

There was a significant main effect of response type on confidence ratings, $F(1, 778) = 116.89$, $p < .001$. Unsurprisingly, target identifications ($M = 78.41\%$, $SE = .92$) were characterized by higher confidence than foil identifications ($M = 58.68\%$, $SE = 1.58$). Further, we investigated whether confidence ratings for target identifications differed based on transfer-appropriate conditions compared with transfer-inappropriate conditions. Here we found a main effect of transfer-appropriate condition, $F(1, 580) = 6.74$, $p = .01$, such that participants were more confident in target identifications made in transfer-appropriate conditions ($M = 80.54\%$, $SE = 1.12$) than in transfer-inappropriate conditions ($M = 75.96\%$, $SE = 1.39$). For foil identifications, this effect was not significant, $F < 1$.

Response latencies. We conducted separate 2 (Encoding Conditions) × 2 (Retrieval Conditions) LMMs for response latencies of target and foil identifications. There was a significant interaction for *target identifications*, $F(1, 578) = 18.21$, $p < .001$. When a partial face was encoded, participants were quicker to identify the target from a partial-face lineup ($M = 5.14$ s, $SE = .23$) than from a full-face lineup ($M = 7.86$ s, $SE = .48$). In contrast, when a full face was encoded, participants took longer to identify the target from a partial-face lineup ($M = 6.47$ s, $SE = .34$) than from a full-face lineup ($M = 6.18$ s, $SE = .35$). There were no other significant effects for response latencies associated with target identifications. There were no statistically significant effects of encoding and retrieval on response latency for foil identifications, $F\text{s} < 1$, $p\text{s} > .33$.

Finally, we investigated the effect of response type on response latency, and we found that target identifications were associated with faster response times ($M = 6.35$ s, $SE = .21$) than foil identifications ($M = 9.28$ s, $SE = .36$), $F(1, 778) = 48.35$, $p < .001$.

Table 1
Results in Experiments 1 (Target-Present Lineups)

Conditions	Target IDs	Foil IDs	Target ID confidence rating	Target ID response latency	d'
Encode full face					
Partial-face lineup	.71 (.03)	.29 (.03)	74.26 (1.97)	6.47 (.34)	1.12
Full-face lineup	**.81 (.03)**	**.19 (.03)**	**79.91 (1.59)**	**6.18 (.35)**	**1.52**
Encode partial face					
Partial-face lineup	**.79 (.03)**	**.21 (.03)**	**81.18 (1.57)**	**5.14 (.23)**	**1.43**
Full-face lineup	.68 (.03)	.32 (.03)	77.73 (1.95)	7.86 (.48)	1.02

Note. Bold rows indicate lineup conditions that match the encoding condition (i.e., transfer-appropriate lineup). Values in parentheses represent standard errors of the mean. Participants were forced to choose a person from the lineup in Experiment 1 ($N = 195$).

Experiment 2

In Experiment 2, we attempted to replicate and extend the novel results of Experiment 1 with masked faces rather than partial, cropped faces. We opted to use masked faces instead of cropped faces to better align our stimuli with how perpetrators mask their faces in real life, thereby increasing ecological validity. The masked faces were generated by superimposing a ski mask onto the full faces. All other aspects were identical to Experiment 1.

Method

Participants and design. The results from Experiment 1 indicated that the effect size of the interaction ($w = .24$) was closer to a medium effect size ($w = .30$) rather than a small effect size ($w = .10$). Therefore, 88 participants were needed to obtain .80 power for a medium effect. To fit the sample size to our counterbalancing scheme (i.e., multiples of 24 participants), we opted to test a minimum of 96 participants in the remaining experiments. Participants were 110 undergraduate Iowa State University students. Data from three participants were excluded from the analysis: two due to a program error and one due to the participant not following instructions. Data from the remaining 107 participants were included in the analysis.[1]

Materials and procedure. The ski mask from Faces™ 4.0 was superimposed over the 12 faces from Experiment 1, leaving only the eyes visible (see Figure 1). Each full-face picture had a resolution of 450 × 350 pixels, and each masked-face picture had a resolution of 543 × 350 pixels. The masked-face pictures were slightly taller than the full-face pictures because the ski mask needed to look like it was being worn. Most importantly, the size of the eyes was identical in the full-face and masked-face versions of the pictures. Again, target stimuli at study and test were the same image (either masked or unmasked). All other materials and procedure were identical to Experiment 1.

Results

Target identifications. The results from Experiment 2 were similar to those from Experiment 1. Most importantly, there was a crossover interaction between encoding and retrieval conditions, $\chi^2(1, 106) = 15.59$, $p < .001$, $OR = 7.70$, CI [2.80, 21.23] (see Table 2). When participants had seen a full face, they identified the target more often in a full-face lineup ($M = .88$, $SE = .03$, $d' = 1.90$) than in a masked-face lineup ($M = .74$, $SE = .04$, $d' = 1.23$). However, when participants had seen a masked face, they identified the target less often in a full-face lineup ($M = .65$, $SE = .05$, $d' = .93$) than in a masked-face lineup ($M = .85$, $SE = .04$, $d' = 1.72$). The main effects for encoding and retrieval were not significant, χ^2s < 1.59, ps $> .21$.

Confidence. For confidence associated with target identifications, encoding and retrieval conditions exhibited a crossover interaction, $F(1, 328) = 6.56$, $p = .01$, (see Table 2). When participants saw a masked face, they made more confident target identifications ($M = 80.56\%$, $SE = 2.24$) in a masked-face lineup than in a full-face lineup ($M = 76.20\%$, $SE = 2.76$). In contrast, when participants saw a full face, they made less confident target identifications ($M = 81.42\%$, $SE = 2.05$) in a full-face lineup than in a masked-face lineup ($M = 73.56\%$, $SE = 2.55$). Confidence ratings for foil identifications, however, did not differ based on encoding or retrieval conditions, all Fs < 2.23, ps $> .14$. Participants were more confident when they identified a target ($M = 78.23\%$, $SE = 1.26$) than when they identified a foil ($M = 59.67\%$, $SE = 2.37$), $F(1, 424) = 47.89$, $p < .01$. Finally, participants were more confident in their target identifications made in transfer-appropriate conditions ($M = 80.99\%$, $SE = 1.51$) compared with transfer-inappropriate conditions ($M = 74.79\%$, $SE = 1.87$), $F(1, 330) = 6.81$, $p = .01$. This effect was not significant for foil identifications, $F < 1$.

Response latencies. For response latencies associated with target identifications, encoding and retrieval conditions exhibited a crossover interaction, $F(1, 330) = 10.40$, $p = .001$. Specifically, when participants saw a masked face, they identified the target more quickly from a masked-face lineup ($M = 5.94$ s, $SE = .41$) than from a full-face lineup ($M = 7.61$ s, $SE = .47$). In contrast, when they saw a full face, they identified the target more slowly from a masked-face lineup ($M = 7.37$ s, $SE = .55$) than from a full-face lineup ($M = 6.13$ s, $SE = .38$). Encoding and retrieval conditions, however, produced no significant effects for the response latencies of foil identifications, Fs < 1, ps $> .39$. Finally, there was a significant effect of response type on response latency, $F(1, 426) = 22.69$, $p < .001$, such that target identifications were associated with faster response times ($M = 6.68$ s, $SE = .24$) than foil identifications ($M = 9.09$ s, $SE = .45$).

Experiment 3

The primary goal of Experiment 2 was to extend the findings of Experiment 1 with the more ecologically valid materials of masked faces as opposed to partial faces. Similar to Experiment 1, the data exhibited a transfer-appropriate pattern for target identification rates, discriminability, confidence, and response latencies. Although these results are promising from the perspective of applying transfer-appropriate processing to lineup construction, so far we have only investigated situations in which participants were forced to choose from a target-present lineup.

It has long been clear that generalizing to eyewitness identification situations requires that eyewitnesses be tested using both target-present and target-absent lineup conditions with a not-present option (e.g., Wells, 1993). Therefore, in Experiment 3, we examined identification performance for both target-absent and target-present lineups. In addition, participants were given the option to respond "not present" from all lineups.

Once again, we predicted that matching lineup modality to encoding would increase identification accuracy for the target-present lineup even when participants were allowed to select no one. However, predictions regarding the influence of transfer-appropriate processing on target-absent lineups are not as straightforward. Little prior research has directly examined the impact of transfer-appropriate processing on false alarms or correct rejections (e.g., many studies have reported recognition accuracy based

[1] One may wonder why we collected more data than planned in Experiments 2–4. This occurred because our experimental signup slots were posted in blocks of eight and 1 week in advance. We posted more slots than we needed because we did not know how many participants would sign up on a weekly basis, and we did not want to cancel the participants who have already signed up for the experiment.

Table 2
Results in Experiments 2 (Target-Present Lineups)

Conditions	Target IDs	Foil IDs	Target ID confidence rating	Target ID response latency	*d′*
Encode full face					
Masked-face lineup	.74 (.04)	.26 (.04)	73.56 (2.55)	7.37 (.55)	1.23
Full-face lineup	**.88 (.03)**	**.12 (.03)**	**81.42 (2.05)**	**6.13 (.38)**	**1.90**
Encode masked face					
Masked-face lineup	**.85 (.04)**	**.15 (.03)**	**80.56 (2.24)**	**5.94 (.41)**	**1.72**
Full-face lineup	.65 (.05)	.35 (.05)	76.20 (2.76)	7.61 (.47)	.93

Note. Bold rows indicate lineup conditions that match the encoding condition (i.e., transfer-appropriate lineup). Values in parentheses represent standard errors of the mean. Participants were forced to choose a person from the lineup in Experiment 2 ($N = 107$).

on hits-false alarms or d'). Among the studies that have reported recognition performance specifically in target-absent situations, the findings are mixed from a transfer-appropriate processing perspective. For example, in lineup identification studies, reinstating the encoding context (e.g., by having participants recall the events leading up to encoding of the target) during retrieval did not affect target-absent lineup performance (i.e., no increase in correct rejections) compared with not reinstating the context (e.g., Krafka & Penrod, 1985). In contrast, studies in word recognition (Hockley, Bancroft, & Bryant, 2012; Murnane & Phelps, 1994) and face recognition (Gruppuso, Lindsay, & Masson, 2007) have shown that reinstating environmental context can sometimes actually increase false alarms (i.e., reduce correct rejections). In light of these findings, we did not make an explicit prediction regarding the effects of matching encoding and retrieval conditions on target-absent lineup performance.

Method

Participants and design. A 2 (Encode: masked face, full face) × 2 (Retrieve: masked-face lineup, full-face lineup) × 2 (Target: absent, present) mixed design was implemented. Unlike Experiments 1 and 2, encoding condition was manipulated between-subjects, whereas retrieval condition and target absence/presence were manipulated within-subjects. Therefore, participants encoded either only full faces or only masked faces, but never both. All participants completed one target-present full-face lineup, one target-absent full-face lineup, one target-present masked-face lineup, and one target-absent masked-face lineup. We tested 200 Iowa State University undergraduate students, but two did not follow instructions and their data were excluded. Ultimately, 98 participants were in the encode masked-face condition and 100 were in the encode full-face condition.

Materials and procedure. To create a three-person target-absent lineup (i.e., replacing the target from the target-present lineup with a foil), an additional four faces were developed using the same parameters as those in Experiments 1. All experimental protocols were identical to Experiment 2 except that each participant encoded only full faces or only masked faces and they were given the option to respond "not present" (Question at retrieval read: "Which pair of eyes did you see or are the eyes you saw not present?"). During the lineup identification task, three full faces or three masked faces were presented horizontally at the center of the screen, with the "not present" option shown centered and below the lineup.

Results

All analyses were done in a similar fashion to those in Experiments 1 and 2, except that we also included signal detection estimates because we have now included both a target-present and a target-absent lineup. Details regarding the signal detection analysis are presented below.

Data from the target-present lineups.

Target identifications. Separate analyses were conducted for the target-absent and target-present conditions. Identification probability was analyzed using GEE with encoding (between-subjects: masked faces, full faces) and retrieval conditions (within-subjects: masked-face lineups, full-face lineups) as the independent variables. Importantly, a crossover interaction was again observed, $\chi^2(1, 197) = 7.83$, $p = .005$, $OR = 3.43$, 95% CI [1.45, 8.12], with no main effects for encoding or retrieval conditions (χ^2s < 1; see Table 3). When participants encoded a full face, they were more likely to identify the target in the full-face lineup ($M = .70$, $SE = .05$) than in the masked-face lineup ($M = .54$, $SE = .05$). However, when participants encoded a masked face, they were less likely to identify the target from the full-face lineup ($M = .59$, $SE = .05$) than from the masked-face lineup ($M = .71$, $SE = .05$). Note that unlike Experiments 1 and 2, we obtained this crossover interaction in Experiment 3 despite allowing participants to choose no one in the lineup.

Foil identifications. Similar to Experiments 1 and 2, foil identifications were not affected by encoding and retrieval conditions, all χ^2s < 1 (see Table 3 for means).

Incorrect rejections. Mirroring the target identification data, a significant crossover interaction was observed for *incorrect rejections* (i.e., choosing "not present" in a target-present lineup), $\chi^2(1, 197) = 13.79$, $p < .001$ (see Table 3). When participants encoded a full face, they were less likely to choose no one in a full-face lineup ($M = .08$, $SE = .03$) than in a masked-face lineup ($M = .28$, $SE = .05$). In contrast, when participants encoded a masked face, they were less likely to choose no one in a masked-face lineup ($M = .07$, $SE = .03$) than in a full-face lineup ($M = .17$, $SE = .04$). The main effects of encoding and retrieval were not significant, χ^2s < 1.49, $ps > .46$.

Confidence. To keep analyses consistent across experiments, we first report results of separate 2 (Encoding) × 2 (Retrieval) LMMs for correct identifications, foil identifications, and incorrect rejections, respectively. We then report the effects of response type on confidence ratings in a one-way LMM.

Table 3
Results in Experiment 3 (Target-Present and Target-Absent Lineups)

Conditions	Correct response	Foil IDs	Incorrect rejections	Correct response confidence	Correct response latency	d'	c
			Target-present lineup				
Encode full face							
Masked-face lineup	.54 (.05)	.18 (.04)	.28 (.05)	74.57 (3.26)	7.09 (.64)	1.15	−.02
Full-face lineup	**.70 (.05)**	**.22 (.04)**	**.08 (.03)**	**77.19 (2.52)**	**7.24 (.71)**	**1.47**	**−.44**
Encode masked face							
Masked-face lineup	**.71 (.05)**	**.21 (.04)**	**.07 (.03)**	**74.91 (2.28)**	**7.46 (.68)**	**1.57**	**−.27**
Full-face lineup	.59 (.05)	.23 (.04)	.17 (.04)	65.24 (2.97)	10.65 (1.18)	.94	−.60
			Target-absent lineup				
Encode full face							
Masked-face lineup	.49 (.05)	.51 (.05)	—	74.90 (3.67)	8.89 (.73)		
Full-face lineup	**.37 (.05)**	**.63 (.05)**	—	**74.46 (3.93)**	**7.81 (.79)**		
Encode masked face							
Masked-face lineup	**.43 (.05)**	**.57 (.05)**	—	**72.67 (3.67)**	**9.37 (.99)**		
Full-face lineup	.23 (.04)	.77 (.04)	—	68.35 (4.78)	10.43 (2.38)		

Note. Correct response refers to Target IDs for target-present lineups and correct rejections for target-absent lineups. Bold rows indicate lineup conditions that match the encoding condition (i.e., transfer-appropriate lineup). Values in parentheses represent standard errors of the mean. Participants were offered a "not present" option. For the encode masked-face condition, $N = 98$, for the encode full-face condition, $N = 100$.

First, the effects of encoding and retrieval conditions on confidence ratings for *target identifications* were examined. A significant interaction was observed, $F(1, 248) = 5.05$, $p = .03$. When participants had encoded a masked face, they were more confident when identifying the target from a masked-face lineup ($M = 74.91\%$, $SE = 2.28$) compared with a full-face lineup ($M = 65.24\%$, $SE = 2.97$), but the opposite occurred when participants had encoded a full face ($M_{masked\text{-}face\text{-}lineup} = 74.57\%$, $SE = 3.26$ vs. $M_{full\text{-}face\text{-}lineup} = 77.19\%$, $SE = 2.52$). In addition, there was a main effect of encoding on target identification confidence, $F(1, 248) = 4.50$, $p = .04$. Specifically, participants made more confident target identifications when they had encoded a full face ($M = 76.05\%$, $SE = 2.00$) than when they had encoded a masked face ($M = 70.53\%$, $SE = 1.88$). The main effect of retrieval was not significant, $F(1, 248) = 1.67$, $p = .20$.

There was a marginal main effect of retrieval condition for *foil identification confidence*, $F(1, 80) = 3.07$, $p = .08$, such that participants were more confident when identifying a foil from a full-face lineup ($M = 65.89\%$, $SE = 3.23$) than from a masked-face lineup ($M = 56.82\%$, $SE = 4.34$). No significant effects of encoding condition and interaction were observed, $F\text{s} < 1.48$, $p > .23$. The confidence ratings for *incorrect rejections* was not reliably affected encoding and retrieval conditions, all $F\text{s} < 1$, $p\text{s} > .32$.

We also examined whether confidence ratings differed based on response type (target identification, foil identification, incorrect rejections). Similar to Experiments 1 and 2, a main effect of response type was observed, $F(3, 393) = 11.13$, $p < .01$, such that participants exhibited greater confidence for target identifications ($M = 73.25\%$, $SE = 1.38$) than foil identifications ($M = 61.68\%$, $SE = 2.67$) and incorrect rejections ($M = 62.28\%$, $SE = 3.22$).

Finally, we examined whether confidence for target identifications made in target-present lineups differed based on transfer-appropriate conditions compared with transfer-inappropriate conditions. Here we found that participants were more confident in target identifications made when they were in transfer-appropriate conditions ($M = 76.05\%$, $SE = 1.70$) compared to when they were in transfer-inappropriate conditions ($M = 69.74\%$, $SE = 2.23$), $F(1, 250) = 5.25$, $p = .02$. This effect was not reliable for foil identifications or incorrect rejections, $F\text{s} < 1.2$, $p\text{s} > .27$.

Response latencies. Encoding and retrieval conditions produced a marginally significant interaction for *target identification* response latencies, $F(1, 248) = 3.41$, $p = .07$. After encoding a masked face, target identifications from a masked-face lineup was associated with faster decision times ($M = 7.46$ s, $SE = .68$) than those made from a full-face lineup ($M = 10.65$ s, $SE = 1.18$). However, when participants encoded a full face, decision times were similar regardless of whether participants made the correct identification from a masked-face lineup ($M = 7.09$ s, $SE = .64$) or a full-face lineup ($M = 7.24$ s, $SE = .71$). In addition, both the main effects of encoding and retrieval conditions were significant, $F\text{s} > 4.07$, $p\text{s} < .05$. Specifically, participants made faster target identifications when they encoded a full face ($M = 7.18$ s, $SE = .49$) than when they encoded a masked face ($M = 8.90$ s, $SE = .66$), and they made faster decisions when identifying a target from a masked-face lineup ($M = 7.30$ s, $SE = .47$) than from a full-face lineup ($M = 8.78$ s, $SE = .69$).

For foil identifications, there was a marginally significant main effect of encoding $F(1, 80) = 3.62$, $p = .06$, and retrieval, $F(1, 80) = 3.25$, $p = .08$, but these variables did not interact, $F(1, 80) = .14$, $p = .71$. Participants were faster to identify a foil when they encoded a masked face ($M = 8.18$ s, $SE = .91$) than when they encoded a full face ($M = 11.20$ s, $SE = 1.42$), and they were faster to identify a foil from a full-face lineup ($M = 8.29$ s, $SE = .89$) than from a masked-face lineup ($M = 11.15$ s, $SE = 1.46$). For incorrect rejections, encoding and retrieval conditions did not affect the response latency, all $F\text{s} < 1$, $p\text{s} > .38$.

Lastly, we found that participants were faster when they made target identifications ($M = 8.05$ s, $SE = .42$) than foil identifications ($M = 9.62$ s, $SE = .84$) and incorrect rejections ($M = 12.99$ s, $SE = 1.26$), $F(2, 393) = 10.99$, $p < .001$.

Target-absent lineups.

Correct rejections. Correct rejection rates in target-absent lineups were analyzed using GEE, with encoding and retrieval conditions as the independent variables. A significant main effect of encoding condition was found, $\chi^2(1, 197) = 4.23$, $p = .04$, $OR = .52$, CI [.28, .97], which indicated that participants correctly rejected the lineup more often when they encoded a full-face ($M = .43$) than when they encoded a masked-face ($M = .33$). The main effect of retrieval condition was also significant. Here, participants were more likely to reject a masked-face lineup ($M = .46$) than a full-face lineup ($M = .30$), $\chi^2(1, 197) = 11.17$, $p = .001$, $OR = 1.64$, CI [.98, 2.74]. However, unlike the data from the target-present lineup, encoding and retrieval conditions did not interact, $\chi^2 < 1$. An examination of Table 3 shows that participants were more likely to reject the masked-face lineup ($M = .43$, $SE = .05$) than the full-face lineup ($M = .23$, $SE = .04$) when they had studied a masked face. Somewhat surprisingly, they were also more likely to reject the masked-face lineup ($M = .49$, $SE = .05$) than the full-face lineup ($M = .37$, $SE = .05$) when they had studied a full face. This pattern of results may indicate that participants used a more conservative response criterion when they encountered a masked-face lineup relative to a full-face lineup. That is, when participants studied a full face, they may have perceived the lineup task to be more difficult when they were shown a masked-face lineup compared with when they were shown a full-face. As a result, participants made more conservative responses. We further consider the feasibility of this account in the General Discussion section.

Confidence. For confidence associated with *correct rejections*, the main effect of encoding condition, retrieval condition, or the interaction was not significant, $Fs < 1.03$, $ps > .31$. However, there was a marginally significant interaction between encoding and retrieval on the confidence ratings of foil identifications, $F(1, 241) = 3.23$, $p = .07$. When participants encoded a masked face, they were more confident in their foil identifications made from a masked-face lineup ($M = 71.43\%$, $SE = 2.66$) than from a full-face lineup ($M = 67.17\%$, $SE = 2.31$). In contrast, when participants encoded a full face, they were less confident in their foil identifications made from masked-face lineup ($M = 62.73\%$, $SE = 3.86$) than from a full-face lineup ($M = 68.68\%$, $SE = 2.65$). The main effects of encoding and retrieval were not statistically significant, $Fs < 1.60$, $ps > .21$. In addition, we found a main effect of response type on confidence ratings, $F(1, 394) = 5.51$, $p = .02$, such that correct rejections were associated with greater confidence ($M = 73.17\%$, $SE = 1.97$) than foil identifications ($M = 67.61\%$, $SE = 1.41$).

Response latencies. Response latencies for *correct rejections* were not affected by the encoding and retrieval variables, $Fs < 1.81$, $ps > .18$. There were also no significant effects of encoding and retrieval conditions on response latency for foil identifications, $Fs < 1.75$, $ps > .19$. Lastly, there was no effect of response type on TA lineup response latency, $F(1, 394) = 1.13$, $p = .29$, with participants making correct rejections ($M = 9.01$ s, $SE = .56$) and foil identifications ($M = 9.76$ s, $SE = .43$) at similar speed in the target-absent lineups.

Compound signal detection model. To examine whether the encoding and retrieval conditions affected discriminability, we employed a compound signal detection model (SDT-CD; see Palmer & Brewer, 2012; Palmer, Brewer, & Weber, 2010) using the integration decision rule when computing discrimination (d') and response bias (c). The integration decision rule posits that decisions are based on a global assessment of the array (i.e., a comparison of the stimulus array against the decision criterion) as opposed to an independent decision rule whereby each stimulus is individually compared with the decision criterion (Palmer & Brewer, 2012). We report estimates based on the integration decision rule because it provided the best fit for the current data ($G^2s < 3.11$, $ps > .08$), a typical procedure when dealing with eyewitness identification data (Palmer et al., 2012).

The CD-SDT model is designed to explain performance for compound decisions which consist of both detection (i.e., whether the target is present among other stimuli) and identification (i.e., whether the perpetrator is in the lineup) components. SDT-CD takes into account all possible response probabilities (target identifications, foil identifications, and incorrect rejections for target-present lineups; correct rejections and foil identifications for target-absent lineups) to find a single d' and response bias (c). The value of response bias ranges from conservative (i.e., negative) to liberal (i.e., positive).

Results for d' and c showed that when participants encoded a masked face, discriminability was lower when they chose from the full-face lineup ($d' = .94$, $c = -.60$) than from the masked-face lineup ($d' = 1.57$, $c = -.27$; $G^2(1) = 14.43$, $p < .01$; see Table 3). Conversely, when participants encoded a full face, discriminability was numerically higher for the full-face lineup ($d' = 1.47$, $c = -.44$) than the masked-face lineup ($d' = 1.15$, $c = -.02$), but this difference did not reach significance, $G^2(1) = 2.27$, $p = .13$.

Experiment 4

In Experiment 3, we found that matching lineup modality to encoding modality increased target identifications, even when participants were offered "not present" as a response option. We have now observed the hypothesized transfer-appropriate pattern in target-present lineups across three experiments. However, the transfer-appropriate pattern was not always obtained when the target was absent from the lineup. Rather, in target-absent conditions participants were more likely to correctly reject the lineup when it featured masked faces rather than full faces, regardless of whether they had encoded a masked or a full face.

For Experiment 4, we aimed to examine why it is more difficult for participants to identify a masked person in a full-face lineup than in a masked-face lineup. Here we consider two possibilities. One possibility is that attempting to identify the eyes in the context of a full face is made difficult by the presence of the noneye features. Specifically, inclusion of these irrelevant features is likely to trigger holistic processing of the face, which may interfere with participants' perception and identification of the eyes (i.e., whole-face interference, Leder & Carbon, 2005). According to this explanation, presenting a full-face lineup to a participant who has encoded a masked face would impair identification performance because the full faces invite holistic processing, leading to a *qualitative* processing mismatch between encoding and retrieval. For exposition purposes, we refer to this as the *processing mismatch account*.

Another potential explanation for the results observed thus far appeals to the idea the difference is *quantitative* rather than qualitative. We begin by focusing on the condition when participants had encoded a masked face. Here, attempting to identify the eyes in a full-face lineup is difficult because inclusion of the irrelevant features increases the "noise" in the visual signal. In the full-face lineup, in addition to encountering three different sets of eyes, participants also saw three different sets of noses, brows, and lips. In contrast, in a masked-face lineup, participants saw three different sets of eyes, but the remaining "features" were held constant by the mask. Therefore, target identification might be more difficult in the full-face lineup than in the masked-face lineup because of the presentation of *more unique features* for which participants had to discount in the former. This situation is not applicable when participants had encoded a full-face. Here, a full-face lineup would facilitate identification because all the facial features in the target would match the encoded face. To facilitate explication, we refer to this as the *feature overload account*.

In Experiment 4, we attempted to tease apart these accounts by holding all irrelevant features in the filler/target faces constant in the full-face lineup (see Figure 1). If the full-face lineup impairs identification of the eyes because it increases the number of unique features in the lineup, then its disadvantage should disappear when the number of the irrelevant noneye features are held constant in the full-face lineup. In contrast, if the full-face lineup impairs identification of the eyes because it encourages holistic, rather than featural, processing of the faces, then its disadvantage should remain even when the irrelevant features of the faces are held constant—because these faces would still be processed holistically, thus leading to a processing mismatch between retrieval and the encoded, featural representation of the masked face.

Method

Participants and design. As in Experiment 3, a 2 (Encoding: masked-face vs. full-face) × 2 (Retrieval: masked-lineup vs. full-lineup) × 2 (Target: absent vs. present) mixed design was implemented. Participants were 226 Iowa State University undergraduates. Three participants did not follow instructions and their data was excluded from analysis, leaving 109 in the encode full-face condition and 114 in the encode masked-face condition.

Materials and procedure. We equated all irrelevant features (i.e., head shape, hair, eyebrows, nose, lips) for the four faces within each category by applying the features of one of the original faces to the remaining three. Each participant encountered two full-face lineups, and the noneye features were different *across* the two full-face lineups, but they were identical *within* each lineup. The materials and procedure were otherwise identical to Experiment 3.

Results

Target-present lineups.

Target identifications. Similar to Experiment 3, the main effects of encoding and retrieval were not significant ($\chi^2 < 1$), but there was a significant crossover interaction between encoding and retrieval, $\chi^2(1, 222) = 14.58$, $p < .001$, $OR = 4.62$, CI [2.11, 10.13] (see Table 4). Specifically, when participants encoded a masked face, they were more likely to identify the target in a masked-face lineup ($M = .75$, $SE = .04$) than in a full-face lineup ($M = .55$, $SE = .05$), but the opposite was true when participants encoded a full face (masked-face lineup $M = .54$, $SE = .05$; full-face lineup $M = .69$, $SE = .05$). Clearly, this result favors the processing mismatch hypothesis rather than the processing overload hypothesis.

Incorrect rejections. Although the main effects of encoding and retrieval conditions were not significant, χ^2s < 1, ps $> .77$, there was a crossover interaction, $\chi^2(1, 222) = 7.27$, $p = .01$ (see Table 4). When participants encoded a full face, they made fewer incorrect rejections in a full-face lineup ($M = .11$, $SE = .03$) than in a masked-face lineup ($M = .21$, $SE = .04$). In contrast, when participants encoded a masked face, they made more incorrect rejections in a full-face lineup ($M = .19$, $SE = .04$) than in a masked-face lineup ($M = .11$, $SE = .03$).

Foil identifications. For foil identifications, there was a significant interaction between encoding and retrieval conditions, $\chi^2(1, 222) = 4.01$, $p = .05$, such that when participants viewed a masked face, they were less likely to identify a foil when they encountered a masked-face lineup ($M = .14$, $SE = .03$) than a full-face lineup ($M = .25$, $SE = .04$). However, when participants saw a full face, they were more likely to identify a foil from a

Table 4
Results in Experiment 4 (Target-Present and Target-Absent Lineups)

Conditions	Correct response	Foil IDs	Incorrect rejections	Correct response confidence	Correct response latency	d'	c
			Target-present lineup				
Encode full face							
Masked-face lineup	.54 (.05)	.25 (.04)	.21 (.04)	65.10 (2.84)	8.26 (.71)	1.03	−.10
Full-face lineup	**.69 (.05)**	**.20 (.04)**	**.11 (.03)**	**70.24 (2.55)**	**6.31 (.47)**	**1.28**	**−.37**
Encode masked face							
Masked-face lineup	**.75 (.04)**	**.14 (.03)**	**.11 (.03)**	**75.97 (2.22)**	**7.70 (.74)**	**1.62**	**−.36**
Full-face lineup	.55 (.05)	.25 (.04)	.19 (.04)	67.78 (2.26)	8.49 (.59)	.98	−.35
			Target-absent lineup				
Encode full face							
Masked-face lineup	.43 (.05)	.57 (.05)	—	72.89 (3.36)	7.67 (.46)		
Full-face lineup	**.31 (.05)**	**.69 (.04)**	—	**66.35 (4.45)**	**10.95 (1.14)**		
Encode masked face							
Masked-face lineup	**.35 (.05)**	**.65 (.04)**	—	**77.25 (3.32)**	**6.89 (.75)**		
Full-face lineup	.35 (.05)	.65 (.04)	—	65.08 (4.17)	8.55 (1.01)		

Note. Correct response refers to Target IDs for target-present lineups and correct rejections for target-absent lineups. Bold rows indicate lineup conditions that match the encoding condition (i.e., transfer-appropriate lineup). Values in parentheses represent standard errors of the mean. Participants were offered a "not present" option. For the encode masked-face condition, $N = 114$, for the encode full-face condition, $N = 109$.

masked-face lineup (M = .25, SE = .04) than from a full-face lineup (M = .20, SE = .04). The main effects of encoding and retrieval were not reliable, χ^2s < 1.16, ps >.28 (see Table 4 for means).

Confidence. For confidence associated with target identifications, there was no main effects of encoding and retrieval, Fs < 2.85, ps > .09, but there was a crossover interaction, $F(1, 279)$ = 7.17, p = .01. When participants encoded a masked face, they were more confident in their target identifications made from a masked-face lineup (M = 75.97%, SE = 2.22) compared to a full-face lineup (M = 67.78%, SE = 2.26). However, when participants encoded a full face, they were less confident in target identifications made from a masked-face lineup (M = 65.10%, SE = 2.84) compared from a full-face lineup (M = 70.24%, SE = 2.55). Moreover, encoding and retrieval conditions did not affect confidence ratings for both foil identifications and incorrect rejections, all Fs < 1.20, ps > .28. In addition, there was a main effect of response type on confidence ratings, $F(2, 443)$ = 4.54, p = .01. Participants were also more confident when making target identifications (M = 70.36%, SE = 1.25) than incorrect rejections (M = 66.35%, SE = 11.25) and foil identifications (M = 55.48%, SE = 2.25).

Further, we examined whether confidence for target identifications made in target-present lineups differed for transfer-appropriate conditions compared with transfer-inappropriate conditions. Here we found that when participants made a target identification in the transfer-appropriate conditions, they were more confident (M = 73.30%, SE = 1.69) compared with when they were in transfer-inappropriate conditions (M = 66.48%, SE = 1.80), $F(1, 281)$ = 7.48, p < .01. This effect was not reliable for foil identifications or incorrect rejections, Fs < 1.

Response latencies. There was a crossover interaction between encoding and retrieval for response latencies associated with target identifications, $F(1, 279)$ = 4.36, p = .04, such that when participants encoded a masked face, they were faster to identify the target from a masked-face lineup (M = 7.70 s, SE = .74) than from a full-face lineup (M = 8.49 s, SE = .59). In contrast, when a full face was encoded, participants were slower to identify the target from a masked-face lineup (M = 8.26 s, SE = .71) than from a full-face lineup (M = 6.31 s, SE = .47). Again, both the main effects of encoding and retrieval conditions were not significant, Fs < 1.53, ps > .22.

Next, we assessed the effect of encoding and retrieval on response latency for foil identifications. The interaction between encoding and retrieval was significant, $F(1, 90)$ = 5.40, p = .02. When participants encoded a masked face, they identified a foil from a masked-face lineup faster (M = 7.50 s, SE = .93) than from a full-face lineup (M = 9.30 s, SE = 1.03). However, when participants originally saw a full face, they were slower to identify a foil from a masked-face lineup (M = 10.72 s, SE = 1.28) than from a full-face lineup (M = 7.47 s, SE = .58). Encoding and retrieval conditions did not affect response latencies of incorrect rejections, Fs < 2.55, ps > .12. Lastly, participants were faster to make target identifications (M = 7.63 s, SE = .33) than foil identifications (M = 8.97 s, SE = .54) and incorrect rejections (M = 9.62 s, SE = .82), $F(2, 443)$ = 4.50, p = .01.

Target-absent lineups.

Correct rejections. Unlike the data of Experiment 3, encoding and retrieval conditions did not influence correct rejections, χ^2s < 1.64, ps > .20. Specifically, when participants encoded a masked face, they were equally likely to reject the lineup when it featured masked faces (M = .35, SE = .05) as when it featured full faces (M = .35, SE = .05). Similar to Experiment 3, however, when participants had studied a full face, they were numerically more likely correctly reject a masked-face lineup (M = .43, SE = .05) than a full-face lineup (M = .31, SE = .05), but this difference did not reach significance, $\chi^2(218)$ = 3.32, p = .09.

Confidence. For confidence ratings associated with correct rejections, the main effect of retrieval was significant, $F(1, 157)$ = 6.02, p = .02. When participants correctly rejected a masked-face lineup they were more confident (M = 74.90%, SE = 2.37) than when they rejected a full-face lineup (M = 65.66%, SE = 3.02). However, the main effect of encoding condition and the interaction were not significant, Fs < 1, ps > .46. For foil identifications, there was a significant main effect of encoding condition, $F(1, 281)$ = 4.04, p = .05. That is, foil identifications were characterized by greater confidence when participants had encoded a masked face (M = 68.11%, SE = 4.32) than when they encoded a full face (M = 58.53%, SE = 2.32). There were no other significant effects, Fs < 2.48, ps > .17. Finally, there was a main effect of response type on confidence ratings, $F(1, 444)$ = 3.85, p = .05, indicating that participants were more confident in when they rejected the target-absent lineup (M = 70.65%, SE = 1.92) than when they identified a foil (M = 63.50%, SE = 2.52).

Response latencies. There was a marginally significant main effect of encoding condition on the response latencies of correct rejections, $F(1, 157)$ = 3.60, p = .06. The main effect of retrieval was also significant, $F(1, 157)$ = .8.65, p = .004. Participants were faster to correctly reject a lineup when they encoded a masked face (M = 7.72 s, SE = .63) compared with when they had encoded a full face (M = 9.05 s, SE = .57), and they were faster to reject a masked-face lineup (M = 7.31 s, SE = .42) than a full-face lineup (M = 9.65 s, SE = .77). The interaction was not significant, F < 1, p = .34.

There was a significant Encoding × Retrieval interaction for the response latencies of foil identifications, $F(1, 281)$ = 5.05, p = .03. When participants encountered a masked face, they were faster to identify a foil from a masked-face lineup (M = 8.14 s, SE = .54) compared with a full-face lineup (M = 10.92 s, SE = .70). However, when participants originally encountered a full face, response time of foil identifications was similar for a masked-face lineup (M = 8.52 s, SE = .57) and a full-face lineup (M = 8.58 s, SE = .58). The main effect of retrieval was also significant, $F(1, 281)$ = 5.49, p = .02, such that participants were faster to identify a foil from a masked-face lineup (M = 8.31 s, SE = .39) than from a full-face lineup (M = 9.74 s, SE = .46). The main effect of encoding was not significant, $F(1, 281)$ = 2.64, p = .11. Lastly, response type did not have a significant impact on response latency for TA lineups, $F(1, 444)$ = 1.68, p = .20 ($M_{correct\text{-}rejection}$ = 8.39 s, SE = .43; $M_{foil\text{-}identification}$ = 9.06 s, SE = .31).

Compound signal detection model. Similar to Experiment 3, results for d' and c indicated that participants showed a transfer-appropriate pattern for discriminability (see Table 4). Specifically, when participants encoded a masked face, discriminability was significantly lower for the full-face lineup (d' = .98, c = −.35) compared with the masked-face lineup (d' = 1.62, c = −.36), $G^2(1)$ = 13.62, p < .01. However, when participants encoded a full face, discriminability was numerically higher when choosing from the full-face lineup ($d' = 1.28$, $c = -.37$) compared with the masked-face lineup ($d' = 1.03$, $c = -.10$), a difference that did not reach significance, $G^2(1) = 2.21$, $p = .14$.

General Discussion

Perpetrators of premeditated crimes often wear masks to conceal their identity. Extant research suggests that disguises such as masks can hinder identification (for a meta-analysis see Shapiro & Penrod, 1986). Here we examined a novel way to construct lineups in the case of a masked perpetrator. Based on the principles of transfer-appropriate processing, we argued that presenting people with a masked-face lineup can enhance their ability to identify a target who was partly concealed by a mask. In four experiments, we found support for this prediction in target-present lineups. Specifically, when participants encoded a partial or masked face, they were more likely to identify the target and showed greater discriminability from a partial or masked-face lineup than from a full-face lineup, regardless of whether they were forced to choose (Experiments 1 and 2) or not (Experiments 3–4) and regardless of whether the irrelevant features in the lineup members were different (Experiment 4) or not (Experiments 1–3). In contrast, when participants encoded a full face, the opposite was true: Participants were better able to identify the target from a full-face lineup compared with a masked-face lineup. Not only did the transfer-appropriate lineups increase target identification rates and discriminability, they also increased confidence ratings, and reduced the response latencies associated with those identifications.

These results serve as preliminary evidence that eyewitnesses may be better equipped to identify a masked perpetrator when the lineup members' faces are masked as opposed to fully exposed. Despite the consistency of our results, we note that one condition is unrealistic. In the real world, when eyewitnesses encounter a perpetrator whose whole face is visible, there is virtually no reason for the investigators to administer a masked-face lineup (i.e., encode full face—retrieve from a masked-face lineup). Nevertheless, we included this condition in the present experiments to thoroughly test the transfer-appropriate processing prediction and to rule out the possibility that masked-face lineups are simply superior to full-face lineups.

Unlike the target-present lineups, the target-absent lineups produced a different pattern of results. Specifically, participants who had viewed a masked face at encoding were more likely to correctly reject the lineup when it consisted of masked faces than full faces. In addition, confidence rating and response latency data followed the correct rejection patterns. This pattern of results is consistent with the transfer-appropriate processing account. However, when participants had encoded a full face, they were again more likely to correctly reject a masked-face lineup than a full-face lineup—a pattern that is inconsistent with the transfer-appropriate processing account. The general superiority of the masked-face lineup for correct rejections is seemingly consistent with a criterion shift account. Specifically, participants might use a more conservative criterion for choosing when they encounter a masked-face lineup, regardless of whether the target was present. We speculate that this more conservative criterion shift may occur when participants were presented with the masked faces, which contained only one feature visible per face. The limited visual cues might cause participants to perceive the masked-face lineup as more difficult than the full-face lineup, which in turn caused them to become more conservative in responding. We caution here that this explanation is post hoc and includes two assumptions. First, for this criterion shift explanation to be valid, it would require participants to switch their response criteria across trials. Second, this criterion shift account assumes that participants become more conservative for masked-face lineups (relative to full-face lineups) only when they have encoded full faces. Specifically, when participants have encoded masked faces, they were less likely to reject a target-present masked-face lineup than a target-present full-face lineup. Regardless of the feasibility of this explanation, for the purpose of implementing transfer-appropriate lineups to improve eyewitness identification, the most important finding from the target-absent lineups is that masked-face lineups produced more correct rejections than the full-face lineups when participants had encoded a masked face. Because it is highly unlikely that one would administer a masked-face lineup to an eyewitness who had seen the full face of the perpetrator.

In Experiment 4, we tested whether the transfer-appropriate pattern observed in target identifications was attributed to a processing mismatch or a feature overload. Under the processing mismatch explanation, the masked-face lineup should facilitate target identification of the eyes because it discourages holistic processing. Therefore, the masked-face lineup superiority should remain even when the irrelevant features of the faces provide no discriminability. In contrast, the feature-overload explanation posits that the masked-face lineup enhances target identification of the eyes because it reduces the number of unique features in the lineup. According to this feature-overload account, the masked-face lineup superiority should be eliminated when only the eyes could be used to discriminate the faces (for a similar idea, see Wixted & Mickes, 2014). The data from Experiment 4 did not support the feature-overload interpretation and instead were clearly consistent with the processing mismatch explanation. This finding suggests that when only a few features are encoded (i.e., under featural processing), target identification suffers when participants are shown a full-face lineup due to the obligatory activation of holistic processing.

Traditionally, researchers treat a perpetrator's use of a mask or disguise as an estimator variable that is not under the control of the legal system (e.g., clarity of view of the perpetrator; Wells, 1978). Therefore, when the perpetrator wears a mask, researchers and investigators alike tend to focus on how it disrupts subsequent identification, with no solution for eliciting more accurate identifications from eyewitnesses who have viewed a disguised perpetrator. For the most part, it is simply noted that identification accuracy suffers when perpetrators were masked compared with when he or she was not (Cutler, Penrod, & Martens, 1987a, 1987b; Mansour et al., 2012; Shapiro & Penrod, 1986; Wells, Memon, & Penrod, 2006). Extensive research efforts have been devoted to examine practices that can enhance eyewitness identification performance for an unmasked perpetrator (e.g., Fitzgerald, Oriet, & Price, 2015; Steblay, Dysart, & Wells, 2011; Wells, 1993; Wells et al., 2006). This contrasts sharply with the dearth of research on appropriate lineup configuration when dealing with a masked perpetrator. Considering the results of the present studies, it might be beneficial to construct a system variable (i.e., a variable that *can* be controlled or manipulated; Wells, 1978) that may be able to mitigate the deleterious effect of disguise on identification accuracy.

Finally, we urge caution in the interpretation of the current results given that the present experiments used several lab-based controls that may not generalize to a real-world crime scenario. Consequently, although we have observed a transfer-appropriate effect in four experiments, we are not yet ready to make procedural recommendations based on our findings for several reasons. First, computer-generated faces may not realistically represent a human face given that they lack the within-face variability of real faces (Burton, 2013). For example, our computer-generated faces do not change its facial expression, and they do not include day to day changes such as signs of fatigue, facial hair, and so forth. Furthermore, one could argue that the present experiments potentially test "picture memory," which involves recognizing identical images between encoding and retrieval (as was the case for most studies in the face recognition literature, Burton, 2013). In contrast, witnesses are often required to identify a perpetrator from a series of mugshots after viewing a live crime, and the appearance of the perpetrator in the mugshot might differ slightly from when the witness first saw the perpetrator during the crime. Future research should aim to examine whether the present transfer appropriate pattern would occur under situations that are more similar to those typical of eyewitness identifications in the real-world (e.g., a mock crime video, a photo lineup).

Second, participant witnesses were instructed to attend to the eyes at encoding for all conditions—an instruction that witnesses would not receive in the real world. We included this instruction to equate encoding intent across conditions. Specifically, in the partial/masked-face encoding condition, the eyes were the only feature available, and participants were thus expected to focus on the eyes. In contrast, when participants were shown a full face, they would likely attend to the whole face and not only the eyes. Therefore, when these participants were shown the eyes during retrieval, they may perform worse than participants who had seen just the eyes during encoding either because they did not focus on the eyes or because of a mismatch in processing. Because our interest was to determine whether a match in *processing* between encoding and retrieval would enhance lineup identification performance, we opted to implement the attend-to-eyes instructions.

Finally, we caution interpretation of the high accuracy rates found in our target-present conditions. In a real lineup, performance may be lower because conditions may be less ideal than they were in the lab. For example, participants were instructed to look at the eyes in all conditions; they received a clear, up-close view of the stimuli; they were exposed to the same stimulus (at encoding and retrieval), as opposed to different viewpoints of the same person; and the retention interval was very short. Thus, less ideal conditions found in real-world crime scenarios may lead to an overall reduction in accuracy rates. It is important that future research also investigate other important variables known to degrade identification accuracy (e.g., encoding time, viewing distance, more variation in faces, etc.; Shapiro & Penrod, 1986).

A common feature of a premeditated crime is that the perpetrator conceals part of his or her face. In these cases, eyewitness identification is one of the most compelling, and sometimes sole, form of evidence (Wells, 1993). Moreover, there are currently no procedural recommendations for how best to deal with masked perpetrator when administrating a lineup. Given the frequency with which crimes are committed under masked disguises, research that is aimed at producing evidence-based procedural recommendations are urgently needed. Here we showed that structuring a transfer-appropriate lineup may improve lineup identification performance for witnesses who have encountered a masked perpetrator.

References

Brewer, N., Weber, N., & Semmler, C. (2013). A role for theory in eyewitness identification research. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology: Vol. II. Memory for people* (pp. 201–218). Mahwah, NJ: Erlbaum.

Bruce, V., & Young, A. (2012). Recognizing faces. In V. Bruce & A. Young (Eds.), *Face perception* (pp. 253–314). New York, NY: Psychology Press.

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 66,* 1467–1485. http://dx.doi.org/10.1080/17470218.2013.800125

Chandler, C. C., & Fisher, R. P. (1996). Retrieval processes and witness memory. In E. L. Bjork & R. A. Bjork (Eds.), *Handbook of perception and cognition* (2nd ed., pp. 493–524). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-012102570-0/50016-1

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987a). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72,* 629–637. http://dx.doi.org/10.1037/0021-9010.72.4.629

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987b). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior, 11,* 233–258. http://dx.doi.org/10.1007/BF01044644

Davies, G., & Flin, R. (1984). The man behind the mask—Disguise and face recognition. *Human Learning: Journal of Practical Research & Applications, 3,* 83–95.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628–634. http://dx.doi.org/10.1037/0096-1523.21.3.628

Fitzgerald, R. J., Oriet, C., & Price, H. L. (2015). Suspect filler similarity in eyewitness lineups: A literature review and a novel methodology. *Law and Human Behavior, 39,* 62–74. http://dx.doi.org/10.1037/lhb0000095

Gruppuso, V., Lindsay, D. S., & Masson, M. E. (2007). I'd know that face anywhere! *Psychonomic Bulletin & Review, 14,* 1085–1089. http://dx.doi.org/10.3758/BF03193095

Hockley, W. E., Bancroft, T. D., & Bryant, E. (2012). Associative and familiarity-based effects of environmental context on memory. *Canadian Journal of Experimental Psychology, 66,* 81–89. http://dx.doi.org/10.1037/a0027136

Hockley, W. E., Hemsworth, D. H., & Consoli, A. (1999). Shades of the mirror effect: Recognition of faces with and without sunglasses. *Memory & Cognition, 27,* 128–138. http://dx.doi.org/10.3758/BF03201219

Innocence Project. (2012a). *New Jersey Supreme Court issues new jury instructions that will greatly improve the way courts handle identification evidence*. Retrieved from https://www.innocenceproject.org/new-jersey-supreme-court-issues-new-jury-instructions-that-will-greatly-improve-the-way-courts-handle-identification-evidence/

Innocence Project. (2012b). *Henderson instructions*. Retrieved from https://www.innocenceproject.org/wp-content/uploads/2017/06/NJ-Jury-Instruction.pdf

IQ Biometrix, Inc. (2003). *FACES, the ultimate composite picture*. Fremont, CA: IQ Biometrix.

Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality and*

Social Psychology, 49, 58–69. http://dx.doi.org/10.1037/0022-3514.49.1.58

Leder, H., & Carbon, C. C. (2005). When context hinders! Learn-test compatibility in face recognition. *The Quarterly Journal of Experimental Psychology, 58,* 235–250. http://dx.doi.org/10.1080/02724980343000936

Luce, R. D. (1963). Detection and recognition. In D. Luce (Ed.), *Handbook of mathematical psychology* (pp. 1–103). Hoboken, NJ: Wiley.

Mansour, J. K., Beaudry, J. L., Bertrand, M. I., Kalmet, N., Melsom, E. I., & Lindsay, R. C. L. (2012). Impact of disguise on identification decisions and confidence with simultaneous and sequential lineups. *Law and Human Behavior, 36,* 513–526. http://dx.doi.org/10.1037/h0093937

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6,* 255–260. http://dx.doi.org/10.1016/S1364-6613(02)01903-4

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533. http://dx.doi.org/10.1016/S0022-5371(77)80016-9

Murnane, K., & Phelps, M. P. (1994). When does a different environmental context make a difference in recognition? A global activation model. *Memory & Cognition, 22,* 584–590. http://dx.doi.org/10.3758/BF03198397

Noble, A. (2013). *Ski masks, a fashion accessory to crime in D.C.: Clearly are problem, but ban not feasible*. Retrieved from http://www.washingtontimes.com/news/2013/jan/30/ski-mask-is-fashion-accessory-to-crime-in-dc/

Palmer, M. A., & Brewer, N. (2012). Sequential lineup presentation promotes less-biased criterion setting but does not improve discriminability. *Law and Human Behavior, 36,* 247–255. http://dx.doi.org/10.1037/h0093923

Palmer, M. A., Brewer, N., & Weber, N. (2010). Postidentification feedback affects subsequent eyewitness identification performance. *Journal of Experimental Psychology, 16,* 387–398. http://dx.doi.org/10.1037/a0021034

People v. Williams, 547 P. 2d 1000, 16 Cal. 3d 663, 128 Cal. Rptr. 888 (C. A. 1976).

Report of the Special Master, State v. Henderson, No. A-8–08. (2011, June 18). Retrieved from http://www.judiciary.state.nj.us/pressrel/HENDERSON%20FINAL%20BRIEF%20.PDF%20%2800621142%29.PDF

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139–156. http://dx.doi.org/10.1037/0033-2909.100.2.139

Simonsohn, U. (2014). *Data colada, thinking about evidence and vice versa: No-way interactions*. Retrieved from http://dx.doi.org/10.15200/winn.142559.90552

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8,* 203–220. http://dx.doi.org/10.3758/BF03196157

State v. Courteau, 461 A. 2d 1358 (R. I. 1983).

State v. Fierro, 489 P. 2d 713, 107 Ariz. 479 (A. Z. 1971).

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. http://dx.doi.org/10.1037/a0021650

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46,* 225–245. http://dx.doi.org/10.1080/14640749308401045

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583–592. http://dx.doi.org/10.3758/BF03211301

Tanaka, J. W., & Simonyi, D. (2016). The "parts and wholes" of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 69,* 1876–1889. http://dx.doi.org/10.1080/17470218.2016.1146780

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471–491. http://dx.doi.org/10.1111/j.2044-8295.1988.tb02747.x

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36,* 1546–1557. http://dx.doi.org/10.1037/0022-3514.36.12.1546

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553–571. http://dx.doi.org/10.1037/0003-066X.48.5.553

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7,* 45–75. http://dx.doi.org/10.1111/j.1529-1006.2006.00027.x

Wilford, M. M., & Wells, G. L. (2010). Does facial processing prioritize change detection?: Change blindness illustrates costs and benefits of holistic processing. *Psychological Science, 21,* 1611–1615. http://dx.doi.org/10.1177/0956797610385952

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review, 121,* 262–276. http://dx.doi.org/10.1037/a0035940

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145. http://dx.doi.org/10.1037/h0027474

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759. http://dx.doi.org/10.1068/p160747

Received March 2, 2018
Revision received August 3, 2018
Accepted August 17, 2018 ■