Author: Mansour, J. K., Beaudry, J. L., Bertrand, M. I., Kalmet, N., Melsom, E. I., Lindsay, R. C. L.
Title: Impact of disguise on identification decisions and confidence with simultaneous and sequential lineups
Year: 2012
Journal: Law and Human Behavior
Volume: 36
Issue: 6
Pages: 513-526
URL: http://doi.org/10.1037/h0093937

Copyright: Copyright © 2012 APA.

This is the author’s version of the work, posted here with the permission of the publisher for your personal use. No further distribution is permitted. You may also be able to access the published version from your library.

The definitive version is available at: http://psycnet.apa.org/
Impact of Disguise on Identification Decision and Confidence with Simultaneous and Sequential Lineups

Jamal K. Mansour
Department of Psychology, Simon Fraser University

Jennifer L. Beaudry
Department of Social Sciences, University of South Carolina Beaufort

Michelle I. Bertrand, Natalie Kalmet, Elisabeth I. Melsom, R. C. L. Lindsay
Department of Psychology, Queen’s University

Author Note
This research was supported in part by grants from the Social Sciences and Humanities Research Council of Canada to Jamal Mansour (767-2007-1734), Jennifer Beaudry (767-2004-1198) and R. C. L. Lindsay (410-09-2674). Portions of this research were presented at the 2007 meeting of the Society of Applied Research in Memory and Cognition. The authors would like to thank Katherine MacLennan, Teresa Mateus, and Jennifer Wood for their hard work in preparing the material for these studies, collecting data, and data analysis; the many students who helped edit the pictures necessary to construct 72 lineups; Marilyn Lindsay for her help at all stages; and Kevin Munhall for his insights about data in general.

Address correspondence to Jamal K. Mansour, Department of Psychology, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6; E-mail: jmansour@sfu.ca
Abstract

In two studies (Ns = 87 and 91) we explored how varying the degree of two types of disguise (stocking mask, or sunglasses and toque [i.e., knitted hat]) affects lineup decisions and confidence from simultaneous and sequential lineups. Correct identifications decreased as the disguise covered more of the mock perpetrator’s face, supporting the conclusion that disguises reduce the amount of information available for encoding. Lineup selections decreased as the proportion of the face disguised increased, but more slowly than accuracy, indicating that witnesses are not attuned to encoding conditions. Correct identifications were most affected by sunglasses ($\eta^2_p = .39$), compared to a toque ($\eta^2_p = .20$) or stocking ($\eta^2_p = .16$). Correct Identifications were similarly low when faces were completely covered with a stocking versus covered to just below the nose, suggesting disguises which disrupt our ability to view faces as a whole may be as detrimental as ones that disrupt specific features. Correct rejections were influenced by sunglasses only ($\eta^2_p = .08$). Lineup type had typical effects on accuracy (i.e., higher correct identifications and lower correct rejections from simultaneous compared to sequential lineups), but did not interact with level of disguise for identification accuracy, and had unexpected effects on confidence.
Impact of Disguise on Identification Decision and Confidence with Simultaneous and Sequential Lineups

Eyewitness identification evidence plays a pivotal role in many criminal cases. Although the police have no control over whether a perpetrator wears a disguise (i.e., it is an estimator variable; Wells, 1978), it clearly negatively impacts identification accuracy (e.g., Shapiro & Penrod, 1986). Thus, the impact of disguise on identification decisions and confidence warrants exploration. In addition, it is critical to understand the mechanisms through which disguise affects witness decisions. Brewer, Weber, and Semmler (2005) identify two ways in which disguise may affect identification accuracy. First, compared to an undisguised perpetrator, viewing a disguised perpetrator provides less identifying information for encoding. As a result, witnesses exposed to a disguised perpetrator must have a less complete memory of the perpetrator. A second, not mutually exclusive, explanation is that disguises may affect witnesses’ perceptions of the difficulty of the identification task. These meta-cognitions may change the decision criterion applied by a witness and subsequently decrease their likelihood of making a selection from a lineup.

A third explanation for the impact of disguise on identification accuracy is encoding specificity—encoding to-be-remembered material in a way that facilitates accuracy at a particular recognition task (Davies & Flin, 1984; McKelvie, 1976; Tulving & Thomson, 1973). As indicated by Shapiro and Penrod's (1986) meta-analysis of face recognition and eyewitness identification studies, altering a person’s appearance between the time of encoding and a face recognition/identification task negatively affects identification accuracy. Davies and Flin (1984) found partial support for this mechanism. They found that correct identifications were highest for faces undisguised at both encoding and recognition, but worst for a face disguised at encoding
but not at recognition. The two intermediate conditions (undisguised at encoding, disguised at test and disguised at encoding and test) did not differ. Patterson and Baddeley (1977) also found partial support for this explanation: identification accuracy varied with whether faces were presented at test in the same way or differently than they were presented at encoding (e.g., with or without a beard). However, if a face was disguised at retrieval, there was no significant difference in accuracy if it was disguised at encoding or not. In summary, encoding specificity may play a role but on its own does not adequately explain the influence of disguise on identification.

Finally, a fourth possibility is that the disguise may influence how a witness allocates their attention. A disguised perpetrator may be interpreted by witnesses as more dangerous. This perception may create additional cognitive load leaving witnesses with fewer processing resources for encoding (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). However, in their meta-analysis looking at the effects of distress on identification, Deffenbacher and colleagues did not find that disguise accounted for significant variance. Similarly, a disguise may distract witnesses by drawing attention to the disguise itself and away from facial information required for later recognition. No research that we are aware of has examined this last possibility.

To date, research has not tested which of these mechanisms—or combination thereof—best accounts for the effects of disguise. Most prior research focuses on how masking particular features influences face recognition (e.g., Sadr, Jarudi, & Sinha, 2003; Terry, 1993), or on the relative impact of disguise compared to other estimator variables (e.g., Cutler, Penrod & Martens, 1987a, 1987b; Shapiro & Penrod, 1986). One goal of the following two experiments was to explore Brewer et al.’s (2005) suggestion that disguise works by decreasing the amount of
information available for encoding by manipulating how much of a to-be-remembered face was covered with a disguise.

Many parts of a face can be disguised, but one that is particularly easy to manipulate is hair. A perpetrator’s hairstyle and facial hair may be disguised and/or changed between a crime (encoding) and a lineup identification (test). Previous research shows that such changes result in a consistent decrease in recognition (Cutler et al., 1987a, 1987b; Patterson & Baddeley, 1977). Moreover, research suggests that the negative impact of disguise on accuracy is specifically related to obstruction of hair cues (e.g., Cutler et al., 1987a, 1987b; Narby, Cutler, & Penrod, 1996; O’Donnell, & Bruce, 2001). In particular, compared to normal exposure, Wright and Sladden (2003) found that viewing targets without hair cues impaired performance in subsequent facial recognition tasks. On the other hand, Yarmey (2004) found that obscuring hair with a baseball cap had no significant effect on accuracy in target-present or -absent lineups. Yarmey notes however, that hair cues may not have been appreciably obscured as his female target’s shoulder-length hair was still visible.

Other research shows that disguising the eyes influences identification accuracy. McKelvie (1976) found that participants made more errors in recognizing faces when eyes were covered at either the encoding or recognition stage. He concluded that eyeglasses seemed to be encoded as part of the eyes. Likewise, the addition of eyeglasses at recognition hinders recognition (Hockley, Hemsworth, & Consoli, 1999; Terry, 1993), possibly because the eyes are a central area of focus (Janik, Wellens, Goldberg, & Dell’Osso, 1978; O’Donnell & Bruce, 2001). Conversely, Patterson and Baddeley (1977) found inconsistent effects of the presence of glasses on the accuracy of facial recognition though they found an overall main effect. Changing the presence or absence of glasses from encoding to test decreased identification accuracy.
compared to no change. Changing from wearing glasses at exposure to not wearing glasses at recognition (and vice versa) in combination with a change in wig, beard, and both had a significant negative effect across pose changes. However, changing from wearing glasses at exposure to not wearing glasses at recognition (and vice versa) in combination with a change of wig had no effect if pose was changed to a profile view at test from a full face view at encoding.

Head coverings and glasses are commonly worn as both accessories and disguises, and, as discussed above, are generally effective as disguises. A second goal of the current experiments is to examine whether one has a more detrimental effect than the other and whether the detrimental effects are additive when these disguise are worn together. Much research on facial recognition indicates that the eyes are the most important facial feature for identifying a face (e.g., Henderson, Williams, & Falk, 2005), though much of this research considers the role of either eyes (e.g., McKelvie, 1976) or hair (e.g., Wright & Sladden, 2003). Additionally, Janik et al. (1978) asked participants which portions of faces they looked at most after being asked to form an impression of each of a series of faces; participants reported looking primarily at the eye and mouth regions. Given that research indicates the eyes are the most important feature for identification and that people report that they use the eyes when reviewing faces, we predicted that masking a target’s eyes would be more detrimental to later identification than masking their hair.

Because hats and sunglasses are worn by many people in everyday life, they are relatively inconspicuous and may not be interpreted as an attempt at disguise. As such, witnesses viewing criminals wearing these accessories may not experience meta-cognitions about the difficulty of identification, feel as threatened by the disguised individual, or as distracted by the disguise. In comparison, more conspicuous disguises (such as a stocking covering the face)
would be expected to induce such meta-cognitions and influence attention, which could in turn
decrease the probability that witnesses would select someone from the lineup. One such
conspicuous disguise is the stocking, which is commonly used and is believed to be effective, but
has received little attention in the literature.

Despite the widespread belief that covering a perpetrator's face with a stocking has a
negative impact on identification accuracy (Bond & McConkey, 1995; van Koppen & Lochun,
1997), we know of only one study that specifically tested this claim. Davies and Flin (1984)
showed that face recognition accuracy (hits) was poorer when targets were studied with a
stocking covering their face compared to without a stocking. In their third experiment, they
demonstrated that the decrease in recognition of targets disguised with a stocking was due to
distortions of facial features, rather than the filtering out of complexion information (i.e., more
global information). At this point, it is unclear whether the impaired identification accuracy they
found was caused by the general flattening of facial features or the flattening of particular
features.

Examining various disguises (i.e., toque, sunglasses, and stocking) is important because
they may have differential effects on recognition (Davies & Flin, 1984). In particular, hats and
sunglasses obscure particular facial features (i.e., hair and eyes), whereas a stocking seems to
disrupt, but not obscure, both the global view of faces and specific features. Recent research
indicates that unfamiliar faces are processed featurally rather than globally (Hancock, Bruce, &
Burton, 2000), suggesting that a stocking should have a lesser effect than a toque and sunglasses
disguise since individual features are more discernible (e.g., compare Figure 1D and Figure 2D).
However, O’Donnell and Bruce (2001) have shown that disruptions to the eyes and mouth
regions, compared to hair and chin regions, have the most detrimental effects on recognition
accuracy. Given that a stocking disguise, but not a toque and sunglasses disguise, disrupts both the eyes and the mouth, one might alternatively predict that the stocking disguise would be more detrimental.

In the current experiments we further explore the effect of a stocking disguise on identification accuracy, and investigate whether identification accuracy and selections generally decrease as disguise covers more of the face. If disguise works by decreasing the amount of information available for encoding, we should see accuracy decrease as more of a target’s face is covered rather than a static negative effect when a face is disguised (to any degree) versus undisguised, as one might expect if meta-cognitions were completely responsible for this effect. We question the impact that disguise will have on witnesses’ meta-cognitions, given previous research suggesting that witnesses are relatively insensitive to viewing conditions—as measured by their identification behavior. For example, Lindsay, Semmler, Weber, Brewer, and Lindsay (2009) found that participant witnesses did not decrease their probability of making a lineup selection as the to-be-identified target was presented further away, even though the accuracy of their selections did decrease.

Additionally, we explore whether a disguise that *distorts* the global view of a face and features (stocking) is less damaging to identification accuracy for unfamiliar faces than a disguise that *obscures* specific features (hat and sunglasses). We expect that feature obstruction will be more detrimental to identification accuracy than distortion for two reasons. First, unfamiliar faces are processed more featurally than globally, and Davies and Flin’s (1984) research suggests that a stocking disguise works by disrupting the encoding of specific features, rather than disrupting the global face pattern. Second, people may expect better performance when they view a target in a toque and/or sunglasses, making them more willing to select from
lineups even when they saw relatively little of the target’s face. An alternative hypothesis is that people will expect to be more accurate when they view faces with a stocking covering because they are still able to see specific features. If this is the case, identification accuracy should be greater for targets viewed in a stocking disguise compared to a hat and sunglasses disguise.

By using the same targets and viewing scenarios across the two types of disguise (i.e., toque and sunglasses versus stocking), we were able to compare how these different disguises impact identification accuracy. We created four levels of disguise for the two different types of disguise. In Experiment 1, participants viewed videos of people with no disguise, wearing a toque (i.e., knitted hat) only, sunglasses only, or both. In Experiment 2, participants viewed videos of people with no disguise, wearing a stocking that covered their hair and forehead, wearing a stocking that covered their head to just below their nose, or wearing a stocking that covered their entire head. Our disguise manipulations enabled comparisons of the four levels of obstruction/disruption to encoding a face, within a disguise, on identification decisions and confidence.

Both experiments also permitted a comparison of accuracy from target-present and target-absent lineups across all disguise conditions and across the two lineup types commonly used in North America: simultaneous and sequential. Simultaneous lineups involve showing all lineup members at once and asking the witness to identify which, if any, of the lineup members perpetrated the crime they witnessed. Sequential lineups involve presenting lineup members one at a time and requiring the witness to make a decision about whether the presented lineup member is the person they saw perpetrate the crime before seeing the next face, without the option to view faces again later and without knowledge of how many lineup members will be presented (Lindsay & Wells, 1985). Neither method is clearly preferable because simultaneous
lineups typically result in more correct identifications while sequential lineups typically result in more correct rejections (Steblay, Dysart, & Wells, 2011).

A number of lineup experiments with disguise manipulations have used either simultaneous or sequential lineups, but only one included both. Cutler and Penrod (1988) exposed participants to targets who had either worn a hat or not, and then presented them with either a simultaneous or sequential lineup. They found marginally significant main effects for type of lineup and for disguise such that identification accuracy was lower when the target wore a hat than when he did not and when simultaneous lineups were used versus sequential lineups. The current experiments extend their findings by using more extensive disguise manipulations and by focusing specifically on the issue of disguise and lineup type. This investigation of the interaction between disguise and lineup type was exploratory in nature, and as such, we did not have any specific hypotheses.

Very little previous research on disguise has not addressed how disguise affects target-absent lineups. Cutler et al. (1987a, 1987b) and Cutler and Penrod (1988) included target-absent lineups but did not compare disguise conditions for target-absent lineups, though they did find disguise significantly influenced accuracy overall (i.e., across target-present and target-absent lineups). O’Rourke, et al., (1989) and Yarmey (2004) also included target-absent lineups but found no effect of disguise on either target-present or target-absent lineups. One might expect that disguise would lead to higher accuracy on target-absent lineups because witnesses would have a fairly salient cue that their memory may not be very detailed (i.e., recollection that the perpetrator wore a disguise). As discussed previously, there is evidence that witnesses do not rely on such meta-cognitions in determining whether to choose from lineups (Lindsay et al., 2009). Given the lack of clear direction from the literature so far, we predict that disguise will have the
same effect on correct rejections as correct identifications: accuracy will decrease with greater disguise because less information is encoded and thus available for making the lineup decision.

Typically, the poorer the viewing conditions, the lower the confidence in lineup decisions (Bothwell, Deffenbacher, & Brigham, 1987). Some research has found that viewing a disguised perpetrator is akin to poor viewing conditions, relative to viewing the same person undisguised (e.g. O’Rourke, Penrod, Cutler, & Stuve, 1989), although other research has not found this relationship (Cutler et al., 1987a; 1987b). Confidence is an important aspect of eyewitness identification because it strongly predicts whether an eyewitness is believed in court and thus has a significant impact on the likelihood of conviction (Cutler, Penrod, & Dexter, 1990). We examine the relationship between disguise and confidence in both experiments.

**Experiment 1**

In Experiment 1 we considered the effect of toque and sunglasses disguises on lineup identifications, lineup rejections, and confidence across simultaneous and sequential lineups. We predicted that the effect of wearing sunglasses would be greater than the effect of wearing a toque, and that combining the toque and sunglasses disguised would have an additive effect. With regard to lineup type, we expected to find the standard pattern of more correct identifications with simultaneous lineups and more correct rejections with sequential lineups. Additionally, we wanted to investigate if and how this pattern would change over levels of disguise. Finally, we expected that as the level of disguise increased, the confidence in identification decisions would decrease.

**Method**
**Participants.** Undergraduate students ($N = 98$; 66 female, $M$ age = 19.08 years, $SD = 2.08$) participated in this experiment for course credit or money. Most participants were of European (.76) or Asian (.16) descent.

**Design.** The design was a 2 (Lineup Type: simultaneous, sequential) x 2 (Target Presence: target-present, target-absent) x 2 (Sunglasses: present, absent) x 2 (Toque: present, absent) mixed design. Lineup type was manipulated between-subjects, whereas target presence, sunglasses, and toque were manipulated within-subjects across 24 repeated trials.\(^1\)

**Materials.**

**Videos.** We created videos of 24 people (12 male, 12 female) of European descent. Each video presented one of four staged scenarios: discussion of a bank robbery, a plot to murder someone, the planning of a burglary with an off-screen accomplice, or the questioning by police after a robbery. Each target acted out each of the four scenarios wearing one of the four possible disguises: no disguise, toque only, sunglasses only, and toque with sunglasses (see Figure 1). Scenarios were counterbalanced such that each one was presented approximately equally across the four disguise conditions. The videos displayed the actors (targets) from the shoulders up. An additional variable, quality of view/data collection date (long, large videos/data collected fall-winter 2006/7 versus short, small videos/data collected fall-winter 2007/8) was manipulated but in the interests of length will not be discussed here. Approximately half of the participants viewed long and large video clips (approximately 30 seconds long, 30 cm high by 23 cm wide),

\(^1\) The experiment was originally designed with 32 targets. Counterbalancing all possible combinations of the 16 male targets and 16 female targets, two scenarios, four disguise conditions, and target-presence/absence would require 1024 conditions, which was deemed unwieldy. Instead, eight sub conditions were created such that for each target, each scenario for each disguise condition was shown across participants. For each participant a factorial combination of target presence, disguise, and scenario was created with a different target being shown in each of the 32 possible conditions (i.e. 2 Target Presence x 2 Toque x 2 Sunglasses x 2 Scenario x 2 Target sex). This yielded 32 trials per participant run in blocks of eight trials. Pilot data indicated that 32 trials led to fatigued participants. Thus, one block of eight trials was dropped, leaving the 24 trials in the experiment reported. Each participant saw every possible combination of disguise (4), by target sex (2) in each block yielding 8 trials per block. Across the three blocks 12 trials were target-present and 12 were target-absent.
whereas the rest viewed short and small video clips (approximately 3 seconds long, and 9 cm high by 6.5 cm wide; when presented on a 43 cm screen). We have no reason to expect a difference in our results as a result of data collection date. This manipulation was designed to increase the heterogeneity of the viewing conditions to ensure variability in response rates and confidence. There were no interactions with this variable for identification accuracy, though there were interactions for confidence. These will be discussed when relevant but interested readers may contact the first author for further information.

**Lineups.** Foils (non-target, known innocent lineup members) for the lineups were selected from a large pool of pictures maintained by the experimental laboratory. Pictures of the foils and targets showed a person (without a disguise) from the shoulders up with a neutral facial expression. Target-present and target-absent lineups were constructed for each target using an iterative matching process (Lindsay & Turtle, 1999). No person appeared in more than one lineup. The individual lineup photos were 5 cm by 7.5 cm when presented on a 43 cm screen, regardless of lineup type. The position of a target in a lineup was either counterbalanced across targets (simultaneous) or randomly selected (sequential), with the targets appearing approximately equally in all six positions, thus avoiding position effects.²

In order to ensure the fairness of our lineups, we recruited two additional sets of participants. The first set of participants \( (N = 30) \) provided descriptions for each target. We randomly selected 12 of these descriptions, and presented them to another independent set of participants \( (N = 36) \), who engaged in a mock witness task. Each mock witness viewed the 24

---
² Due to an error in stimuli development, targets were shown in position one five times and position six only three times for simultaneous lineups. For all other positions targets were shown in each position four times. Following the experiment, we counted the frequency with which targets occupied particular positions in sequential lineups. The distribution of positions was similar to that for simultaneous lineups. That is, across all participants, targets were presented most often in position one and least often in position six, with positions two through five being shown intermediately often. For both lineup types, z-tests indicated that the difference in the frequency with which targets were shown in each position were not significantly different.
target-present lineups in a simultaneous fashion, with each lineup presented with one of 12 possible descriptions of that particular target. Collapsing across descriptions and lineups, Tredoux’s E ranged from 2.12 to 5.23 ($M = 3.89$, $SD = .98$; Malpass, Tredoux, & McQuiston-Surrett, 2007). Given these results, we were satisfied that, in general, our lineups were fair.

**Lineup instructions.** Eyewitness participants read the lineup instructions prior to viewing the set of 24 video-lineup pairs. They were told that the “criminal” (target) may or may not be present in the lineup, that they could make only one selection per lineup, and that once selected, they could not change their answer. In the simultaneous condition, participants were told to select the number corresponding to the “criminal’s” position, or to select *not there* if the “criminal” was not present. In the sequential condition, participants were instructed to select *yes* if the presented picture was of the “criminal” and *no* if the picture was not of the “criminal.”

**Procedure.** Participants were randomly assigned to sequential or simultaneous lineups, with approximately half in each. Within each lineup condition, participants were randomly assigned to one of eight possible sub-conditions that varied with respect to which lineups were target-present or -absent and which videos were shown (varying by scenario, target, and level of disguise) with the stipulation that the number of participants run in each sub-condition remain similar. This assured that all possible stimulus and lineup combinations were used approximately equally often.

Participants sat at a private computer terminal. A letter of information revealed that participants would complete a series of trials where they would watch a brief video and then attempt to identify the person from the video from a subsequent lineup. After signing a consent form, participants entered their sex, age, and ethnicity. Participants then read the lineup instructions and completed 24 trials. For each trial, participants watched a video, made an
identification decision from a lineup, provided a confidence statement in that decision, and answered a prior knowledge question (i.e., asking them if they recognized anyone in the lineup from somewhere other than within the experiment). The 24 trials were divided into three randomly ordered blocks. Within each block, four male and four female targets were presented, with one male and one female appearing in each of the four levels of disguise. For each level of disguise, one lineup was target-present and one was target-absent (e.g., if the male lineup for a particular disguise was target-present, the female lineup for the same disguise was target-absent). Participants had no prior knowledge of the number of target-present and -absent lineups. Between each block of eight, participants engaged in a one minute filler task to prevent fatigue and boredom. At the end of the three blocks, participants were debriefed.

After each video, participants in the simultaneous condition were shown all six lineup members at once. They could select someone as the target by selecting the number corresponding to the target’s position, or indicate that the target was not present by selecting *not there*. Participants viewing sequential lineups were shown up to six lineup members, one at a time, and indicated whether each lineup member was the target by selecting *yes* if the presented photo was of the “criminal” or *no* if it was not. If participants chose *no*, lineup members continued to be displayed until participants either chose *yes* or had viewed all six lineup members. The lineup stopped if participants chose *yes* to a lineup member. Participants could only view each lineup member once and were not informed of how many pictures were in the lineup (though this may have become obvious over the course of the experiment).

After completing each lineup, participants rated their confidence in their identification decision from 0% (not at all confident) to 100% (extremely confident). In the sequential condition, the confidence rating was collected after the participant either selected someone as the
target or indicated *no* to all six lineup members. Participants also were asked if they recognized someone in the lineup from real-life (prior knowledge question). If the participant answered *yes*, they were asked to indicate which lineup member or members were recognized and where the lineup members had been encountered.³

**Measures.**

**Identification Accuracy.** For each participant, we calculated the proportions of correct and incorrect selections and rejections. Selections of targets from target-present lineups were correct identifications. Any selection from a target-absent lineup was an incorrect selection. Saying “not there” to all lineup members (either collectively for simultaneous lineups or individually for sequential lineups) was a rejection. Rejections of target-absent lineups were correct rejections and rejections of target-present lineups were incorrect rejections. We also calculated the proportion of decisions in which participants made any selections (target or foil) from target-present lineups (referred to as target-present selections). For target-absent lineups, the selection rate is simply 1 minus the correct rejection rate.

**Confidence.** Mean confidence was calculated for correct identifications and correct rejections.

**Results**

Witnesses to crimes typically do not review multiple lineups but the participants in our experiment viewed 24 lineups. We used binomial logistic regression with trial as the predictor to determine whether learning accounted for performance on target-present and target-absent lineups. For target-present lineups, selections of the target were coded as correct responses; all

---

³ We included this question because the images collected for foils were photographs of introductory psychology students from the previous year. Moreover a couple of the targets were teaching assistants for the introductory psychology course from which participants were selected. Teaching assistants were not available for the entire term so in many cases, there was no reason to expect targets to be known to participants; however, we wanted to ensure the results were not influenced by prior knowledge.
other responses were coded as incorrect. For target-absent lineups, rejections of the lineup were
coded as correct responses and identification of any lineup member as an incorrect response. No
learning effects were found ($p > .10$).

On average, participants recognized lineup members from outside of this study from 2.56
lineups ($SD = 1.80$; $Range = 1–9$). In order to ensure recognition rates were not inflated by prior
knowledge and to maintain a set of 24 trials per participant, we dropped 11 participants who
indicated that they recognized a target face (versus a lineup foil) from somewhere other than the
experiment. For the remaining participants ($N = 87$), correct identifications, correct rejections,
target-present selections, confidence in correct identifications, and confidence in correct
rejections were examined with $2 \times 2 \times 2$ (Toque: present, absent) $\times$ (Sunglasses: present, absent) $\times$ (Lineup Type: simultaneous, sequential) mixed-model repeated-measures analyses of variance
(ANOVAs).

**Correct identifications.** The correct identification rate was significantly lower when the
target wore sunglasses ($M = .61$, $SD = .24$) than when the target wore no sunglasses ($M = .82$, $SD
= .17$), $F(1, 84) = 53.15$, $p < .001$, $\eta^2_p = .39$. Likewise, correct identifications were significantly
lower when the target was wearing a toque ($M = .66$, $SD = .21$) than when the target was not
wearing a toque ($M = .77$, $SD = .17$), $F(1, 84) = 20.51$, $p < .001$, $\eta^2_p = .20$. Planned comparisons
also showed that the correct identification rate was significantly lower if the targets wore
sunglasses alone than a toque alone, $F(1, 84) = 7.95$, $p = .006$, $\eta^2_p = .09$, and that the effect of
wearing both sunglasses and toque was larger than the effect of either alone, $F(1, 84) = 25.20$, $p$

---

4 The mean for toque is the mean correct identification rate for lineups when the target wore a toque alone or wore
both a toque and sunglasses. The mean for no toque refers to the no disguise and sunglasses only conditions. Means
for sunglasses versus no sunglasses were calculated similarly. In-text means for correct rejections and target-present
selections are presented this way as well.
< .001, \( \eta_p^2 = .23 \) (see Table 1 for means\(^5\)). Both results are consistent with our expectations.

Consistent with previous research there were more correct identifications when participants viewed simultaneous (\( M = .77, SD = .15 \)) compared to sequential lineups (\( M = .66, SD = .15 \)), \( F(1,84) = 9.93, p = .002, \eta_p^2 = .11 \). There were no significant interactions (\( ps > .18 \)).

**Correct rejections.** For correct rejections, participants made fewer correct rejections when the target wore sunglasses (\( M = .65, SD = .23 \)) than when they did not (\( M = .72, SD = .24 \)), \( F(1,84) = 7.14, p = .009, \eta_p^2 = .08 \). The effect of toque was not significant (\( p = .62 \)); participants made just as many correct rejections when the target wore a toque (\( M = .69, SD = .27 \)) as when the target did not wear a toque (\( M = .68, SD = .23 \)). As such, we found that the effect of sunglasses was significantly greater than the effect of toque, \( F(1,84) = 4.18, p = .04, \eta_p^2 = .05 \) and the combined effect of toque and sunglasses was not greater than either individually (\( p = .96 \)). The effect of lineup type was marginally significant with more correct rejections when participants saw sequential (\( M = .73, SD = .28 \)) compared to simultaneous lineups (\( M = .65, SD = .27 \)), \( F(1,84) = 4.69, p = .06, \eta_p^2 = .04 \). Again, there were no significant interactions (\( ps > .18 \)). Mean correct rejections for each disguise condition and lineup type are available in Table 1.

**Target-present selections.** For target-present sunglasses trials, participants made an average of .90 (\( SD = .13 \)) selections (correct identifications and foil selections) when targets did not wear sunglasses compared to .77 (\( SD = .22 \)) when the targets did wear sunglasses, \( F(1,84) = 26.55, p < .001, \eta_p^2 = .22 \). For target-present toque trials, participants made an average of .87 (\( SD = .14 \)) selections when the target had no toque compared to .80 (\( SD = .19 \)) when the target had a toque, \( F(1,95) = 15.78, p < .001, \eta_p^2 = .14 \). The main effect of lineup type on target-present

---

\(^{5}\) Please note that the tables reporting identification rates and confidence for Experiments 1 and 2 provide the means for the conditions run (i.e., no disguise, toque only, sunglasses only, toque and sunglasses), rather than the estimated marginal means for the ANOVAs which look at toque versus no toque and sunglasses versus no sunglasses. This is necessary in order to provide means for all possible responses (i.e., target selections, foil selections, lineup rejections) as very few participants provided all possible types of responses for each disguise condition.
selections was not significant \((p = .10)\), though the selection rate for simultaneous lineups \((M = .85, SD = .14)\) was higher than for sequential lineups \((M = .80, SD = .15)\). There were no significant interactions \((ps > .20)\). Mean rates of target-present selections for disguise conditions and lineup type can be estimated from Table 1.

**Confidence in correct identifications.** Confidence in correct identifications was lower when the target wore sunglasses \((M = 73.12\%, SD = 15.54)\) than when the target did not wear sunglasses \((M = 86.64\%, SD = 11.15)\), \(F(1, 75) = 41.45, p < .001, \eta^2_p = .36\). Likewise, participants’ confidence was lower when the target wore a toque \((M = 56.02\%, SD = 20.98)\) than when the target did not wear a toque \((M = 60.41\%, SD = 20.43)\), \(F(1, 75) = 15.69, p < .001, \eta^2_p = .17\). Overall, confidence was higher for simultaneous \((M = 73.71\%, SD = 20.38)\) than sequential lineups \((M = 42.73\%, SD = 20.30)\), \(F(1, 75) = 44.42, p < .001, \eta^2_p = .37\). These main effects were qualified by a marginally significant interaction of sunglasses with lineup type, \(F(1, 75) = 3.53, p = .06, \eta^2_p = .05\). There was a smaller difference in confidence ratings between no sunglasses and sunglasses for sequential \((5.73\% \text{ mean difference})\) compared to simultaneous lineups \((11.20\% \text{ mean difference})\). There was also a marginally significant interaction of quality of view of the target (i.e., long, large videos versus short, small videos) and toque, \(F(1, 75) = 3.77, p = .06, \eta^2_p = .08\), whereby confidence in correct identifications was lower in the toque than the no toque condition for participants who viewed the longer, larger videos at encoding \((p < .001)\), but not significant for participants who viewed the shorter, smaller videos \((p = .18)\). Table 2 contains mean confidence ratings for each disguise condition and lineup type.

**Confidence in correct rejections.** Confidence in correct rejections was lower when the target wore sunglasses \((M = 49.69\%, SD = 19.19)\) than when the target did not wear sunglasses \((M = 54.70\%, SD = 19.80)\), \(F(1, 71) = 13.24, p = .001, \eta^2_p = .16\). There was no effect of toque \((p\)
Participants were less confident when targets wore a toque ($M = 36.53\%$, $SD = 26.23$) than when they did not ($M = 39.45\%$, $SD = 25.41$) for sequential lineups ($p < .001$). However, participants were more confident when they wore a toque ($M = 67.93\%$, $SD = 28.24$) than when they did not ($M = 64.88\%$, $SD = 27.35$) for simultaneous lineups ($p < .001$). Toque again interacted with quality of view of the target, $F(1,71) = 5.76$, $p = .02$, $\eta^2_p = .08$, such that toque had a marginally significant effect on confidence in correct rejections for participants who had a good view of the target ($p = .06$) but not those who had a poor view of the target ($p = .12$). Overall, participants were more confident in correct rejections of simultaneous ($M = 66.40\%$, $SD = 18.83$) than sequential lineups ($M = 37.99\%$, $SD = 18.49$), $F(1,71) = 42.45$, $p < .001$, $\eta^2_p = .37$ (also see Table 2).

**Discussion**

Experiment 1 replicated previous findings showing that disguise, in this case a toque and sunglasses, reduces identification accuracy. Consistent with our hypotheses, sunglasses had a more detrimental effect on identification accuracy than the toque, regardless of lineup type or target presence. The toque disguise only influenced target-present lineup decisions, suggesting that perhaps we use hair information as a confirmatory cue when we see a face that matches our memory for a perpetrator. Consistent with this explanation, the combined effect of wearing a toque and sunglasses on correct identifications (though not correct rejections) was greater than wearing either alone. The effects of disguise on target-absent lineups were similar to target-present lineups: accuracy did indeed decrease with a target had been disguised; however, the effects were quite a bit smaller than with target-present lineups.
Simultaneous lineups resulted in more correct identifications than sequential lineups, while sequential lineups resulted in more correct rejections than simultaneous lineups. Thus, identification results replicate typical findings with these types of lineups (Steblay et al., 2011).

The findings for confidence were much less clear. For correct identifications, confidence decreased with degree of disguise which is unsurprising and promising—participants demonstrated sensitivity to encoding conditions by decreasing confidence in their identifications. This is consistent with O’Rourke et al. (1989), but contrary to Cutler et al. (1987a; 1987b), suggesting the relationship between confidence and disguise is likely influenced by a third variable, perhaps meta-cognitions. We were surprised to find that when participants viewed a simultaneous lineup, they were more confident in correct rejections if the target had worn a toque than if they had not worn one. This relationship was in the opposite direction for sequential lineups. Previous research examining confidence of correct rejections from simultaneous and sequential lineups has either found no difference (Lindsay & Wells, 1985) or higher confidence for simultaneous lineups (Gronlund, Carlson, Dailey, & Goodsell, 2009).

Perhaps cognitive dissonance is responsible for these unexpected findings for confidence (Festinger & Carlsmith, 1959). Because simultaneous lineups may encourage a more liberal response criterion, participants may have been less willing to reject a lineup than if they were presented with sequential lineups (the significant lineup type main effect supports this explanation). In order to justify the rejection of a lineup (i.e., not selecting any of the lineup members), participants may have made rejections with greater confidence than when in the absence of such a belief. The conservative approach encouraged by sequential lineups could have led participants to reject a lineup if they did not believe they could make an accurate identification and thus, there would be no need to alleviate cognitive dissonance. However, it is
unclear why this effect would have occurred for the toque but not the sunglasses. Notably, confidence levels for correct identifications across simultaneous and sequential lineups follow the same pattern as the correct rejection interaction, but did not reach significance.

In the following experiment we investigate the impact of a second type of disguise, the stocking, on identification accuracy and confidence.

### Experiment 2

Experiment 2 assesses the impact of a stocking, partially or completely covering the head, on lineup identifications. Davies and Flin (1984) used this manipulation and concluded that stockings reduced recognition accuracy because they changed the nature of the features viewed during the encoding versus the recognition task.6 Davies and Flin examined only target-present simultaneous lineups. In the current experiment, we replicate and extend their work using both target-present and target-absent simultaneous and sequential lineups. As well, we examine the effectiveness of a partial disguise. That is, compared to covering the face entirely, how is identification accuracy affected when the target’s face is partially covered with a stocking, thereby allowing some features to be encoded?

### Method

**Participants.** Undergraduate university students \(N = 102; 74\) females, \(M\) age = 18.94 years, \(SD = 1.66\) participated in the experiment. Most participants were European (.70) or Asian (.20). No one who participated in Experiment 1 participated in Experiment 2. Participants received either course credit or money for participating.

**Design.** We used a 2 (Lineup Type: simultaneous, sequential) x 2 (Target Presence: target-present, target-absent) x 4 (Disguise: none, 1/3, 2/3, full [i.e., all of the head covered by a

---

6 It is worth noting here that Davies and Flin (1984) did not use a classic lineup paradigm. They presented participants with a series of four faces and later asked them to choose these four targets from an array of 16 faces.
stocking]) mixed design, with lineup type as a between-subjects factor, and disguise and target presence as the within-subjects factors. The nature of the disguise manipulation is described below.

Experiment 2 involved the same materials, procedures, and measures as Experiment 1 except where indicated.

Materials.

Videos. Videos of the same 24 targets from Experiment 1 served as the stimuli in this experiment. Instead of the toque and sunglasses disguise, the targets wore a diaphanous stocking pulled down from the top of their head. The stocking disguise had four levels: no stocking at all, a stocking covering one third of their face (hair and forehead covered); a stocking covering two-thirds of their face (hair, forehead, eyes, and nose covered); or a stocking covering their entire head (see Figure 2). Each target video depicted one of two scenes, discussion of a bank robbery or the planning of a burglary with an off-screen accomplice. The on-screen video sizes were the same and the video durations were similar to Experiment 1. There were no interactions with quality of view of the target variable on any of the measures.

Lineups. Lineups were the same as those used in Experiment 1, though the specific position of faces (simultaneous lineups) and order of faces (sequential lineups) differed.

Results

We used binomial logistic regression to confirm that no learning effects occurred ($p$s $> .10). Participants indicated prior knowledge of at least one lineup member for an average of 2.60 trials ($SD = 1.33$; $Range: 1 – 7$). Eleven participants who indicated they recognized a target from outside of the task were dropped from analysis. For the remaining 91 participants, we examined the impact of disguise and lineup type on correct identifications, correct rejections, selections
from target-present lineups, confidence in correct identifications, and confidence in correct rejections using 4 (Disguise: none, 1/3, 2/3, fully covered) x 2 (Lineup Type: simultaneous, sequential) mixed-model repeated-measures ANOVAs.

**Correct identifications.** Correct identifications were highest for trials in which the actor wore no disguise, lower when 1/3 of the face was covered, and lowest for a 2/3 covered face and a fully covered face, which did not differ, $F(3, 264) = 17.31, p < .001, \eta_p^2 = .16$ (see Table 3). The expected main effect of lineup type was not found ($p = .76$); correct identifications were approximately equal when participants saw simultaneous lineups ($M = .64, SD = .16$) compared to sequential lineups ($M = .63, SD = .16$). There was no significant interaction ($ps > .65$).

**Correct rejections.** There was no main effect of disguise on correct rejections ($p = .57$; see Table 3). However, the expected main effect of lineup type was present: participants who viewed sequential lineups ($M = .75, SD = .19$) made significantly more correct rejections than participants who viewed simultaneous lineups ($M = .66, SD = .19$), $F(1, 88) = 5.07, p = .03, \eta_p^2 = .05$. The interaction of lineup type with disguise was not significant ($p = .55$).

**Target-present selections.** The target-present selection results mirrored correct identifications: target-present selections decreased from the no disguise condition to the 2/3 and fully covered disguise conditions, which did not differ, $F(3, 264) = 11.36, p < .001, \eta_p^2 = .11$ (see Table 3). The main effect of lineup type was marginally significant, with simultaneous lineups leading to more selections ($M = .83, SD = .15$) than sequential lineups ($M = .77, SD = .15$), $F(1, 88) = 3.15, p = .08, \eta_p^2 = .04$). There was no significant interaction ($ps > .13$).

**Confidence in correct identifications.** Confidence in correct identifications decreased as degree of disguise increased such that confidence was highest when targets wore no disguise and lowest when targets had their face 2/3 or fully covered, which did not differ, $F(3, 189) = 31.31, p$
< .001, $\eta^2_p = .33$ (see Table 2). There was no main effect of lineup type ($p = .51$) and no interaction ($p = .44$).

**Confidence in correct rejections.** There was a main effect of disguise on confidence in correct rejections, $F(3, 216) = 9.40, p < .001, \eta^2_p = .12$. The 2/3 disguise condition led to significantly lower confidence than all other disguise conditions ($ps < .001$) while there was no difference among the other three levels ($ps > .60$; see Table 2). There was no significant main effect of lineup type ($p = .54$), but, as in Experiment 1, there was a disguise by lineup type interaction, $F(3, 216) = 3.04, p = .03, \eta^2_p = .04$. Pairwise comparisons indicated that participants were more confident in correct rejections of simultaneous than sequential lineups if the target was undisguised ($p = .05$), but confidence did not differ for the other three levels of disguise ($ps > .27$; see Table 2).

**Discussion**

As in Experiment 1, and consistent with Davies and Flin (1984), disguise had a significant impact on correct identifications and target-present selections and a similar but non-significant effect on correct rejections. Interestingly, covering the perpetrator’s face to just below the nose significantly reduced correct identifications as much as covering the face completely. This leads us to question the role of the mouth and chin for face recognition. Past research indicates the eyes are most important feature (Janik et al., 1978; McKelvie, 1976; O’Donnell & Bruce, 2001), although research on speech perception (e.g., Everdell, Marsh, Yurick, Munhall, & Paré, 2007) indicates that when watching videos of people speaking (as the participants did in this experiment) the mouth and eyes are fixated upon.

Alternatively, perhaps, even though the face was not entirely obscured, the distracting effect of a salient line across the face distorted the global facial appearance, leading to similar
levels of accuracy between the 2/3 and fully covered disguise conditions for target-present lineups. Certainly the 2/3 covered condition appears to have had a greater effect on confidence for both target-present and target-absent lineups than did fully covering the target’s face. The current results provide further support for Davies and Flin’s (1984) suggestion that a stocking disguise leads to decreased identification accuracy because it distorts facial features. Moreover, this experiment suggests that disrupting the global configuration of a face can be just as detrimental as disrupting features, though in the 2/3 covered condition we surely disrupted features as well as the global appearance.

It is also possible that feature information was disrupted enough in the 2/3 disguise condition that additional disruption of the mouth and chin do not further benefit the disguised individual. This is consistent with previous research by O’Donnell and Bruce (2001) showing that changes to the eyes and hair are more disruptive to facial recognition than changes to the mouth and chin. Further research is needed to understand the role of each feature. Using eye tracking to compare which features are looked at when faces are variously disguised could be a productive avenue of study.

In Experiment 1, sunglasses (but not toque) had a significant impact on correct rejections. In the current experiment, the stocking disguise did not have a significant main effect on correct rejections. However, correct rejections followed the same pattern as correct identifications. Given that participants were very willing to make an identification in general—the overall rate of selections was .53—it seems that witnesses may be unwilling to reject lineups unless the most important information for identification, the eyes, are obscured from view. This suggests a meta-cognitive explanation: witnesses may believe that they should be making an identification (despite unbiased instructions) and fail to take into account the amount or quality of information
of the perpetrator’s face when determining whether to select someone from a lineup, thus leading to a low rate of correct rejections. Nonetheless, after they have made a selection from the lineup, witnesses seem to take quality of information into account when rating their confidence in that decision. Indeed, we found that confidence in correct identifications decreased as targets were increasingly disguised.

However, the influence of disguise on confidence in correct rejections was again very interesting. In both experiments we found an interaction of disguise and lineup type. However, the nature of the interaction differed across experiments. In Experiment 1 participants were more confident in correct rejections when they made decisions about simultaneous versus sequential lineups and we explained this using cognitive dissonance theory. In Experiment 2, confidence in correct rejections was higher for simultaneous than sequential lineups when targets were undisguised but approximately equivalent otherwise. This is at odds with the explanation of cognitive dissonance posed for the results of Experiment 1. That is, we posited that participants expected to be able to select someone from the lineup and when they could not, this created cognitive dissonance which was relieved by being highly confident in their decision.

The discrepancy between results across the two experiments may reflect the nature of the two disguises tested. As discussed earlier, stockings (Experiment 2) are more conspicuous than a toque and sunglasses (Experiment 1). As a result, participants exposed to a stocking disguise may have expected that it would negatively impact their ability to accurately select someone from the lineup. Participants exposed to the toque and sunglasses disguise may not have had the same expectation, given the inconspicuous nature of the disguise. The expectation that a stocking disguise would influence accuracy may have translated into witnesses’ confidence ratings such that when a rejection was made, there was no dissonance to relieve unless the rejection was for a
lineup following an undisguised target (i.e., there was no reason to expect a negative effect of disguise on accuracy).

**Additional Analyses to Compare Experiment 1 and 2**

The data for Experiments 1 and 2 were collected simultaneously (i.e., participants were randomly assigned to participate in either), and thus, it is reasonable to compare the two disguises statistically in order to determine which disguise had a greater impact. In order to do this the 2 Sunglasses x 2 Toque design of Experiment 1 was sorted into four levels of disguise, based on the correct identification rates (i.e., no disguise, toque only, sunglasses only, toque and sunglasses). We then conducted mixed-model repeated-measures ANOVAs with degree of disguise as a within-subjects factor and lineup type (simultaneous, sequential) and type of disguise (toque/sunglasses, stocking) as between-subjects factors on correct identifications, correct rejections, and target-present selections.

Looking first at correct identifications, there was a significant main effect of lineup type whereby there were more correct identifications when participants viewed simultaneous ($M = .71$, $SD = .16$) versus sequential lineups ($M = .66$, $SD = .16$), $F(1, 174) = 5.24$, $p = .02$, $\eta^2_p = .03$. The main effect of disguise was significant, $F(3, 522) = 40.24$, $p < .001$, $\eta^2_p = .20$, confirming that the more disguised a target was, the less likely they were to be correctly identified. There was also a significant main effect of type of disguise, with participants making more correct identifications in the toque and sunglasses experiment ($M = .72$, $SD = .16$; Experiment 1) than in the stocking experiment ($M = .64$, $SD = .16$; Experiment 2), $F(1, 174) = 10.44$, $p = .001$, $\eta^2_p = .06$. We will further discuss the impact of disguise type in the general discussion.

These main effects were qualified by interactions. There was a marginally significant interaction of type of disguise and lineup type, $F(1, 174) = 3.67$, $p = .06$, $\eta^2_p = .02$ whereby the
toque and sunglasses ($M = .77, SD = .16$) experiment led to greater correct identifications than the stocking experiment ($M = .65, SD = .16$) for simultaneous lineups ($p < .001$) but not for sequential lineups ($M = .67, SD = .16$ for toque and sunglasses, $M = .64, SD = .16$ for stocking; $p = .34$). These results emphasize that simultaneous lineups lead to a greater willingness to choose than sequential lineups and indicates that this willingness does vary somewhat with witnessing conditions.

There was also a significant interaction of type of disguise and degree of disguise for correct identifications, $F(3, 174) = 2.86, p = .04, \eta^2_p = .02$. When the target wore no disguise, Experiment 1 had a marginally higher correct identification rate than Experiment 2 ($p = .06$). Given that participants were randomly assigned to Experiment 1 or 2 and the use of identical non-disguised stimuli, this may be a chance finding. More importantly, the toque and sunglasses disguise led to significantly more correct identifications than the stocking in the two intermediate disguise conditions (toque only vs. 1/3 covered, and sunglasses only vs. 2/3 covered; $p_s < .05$). There was no difference in correct identifications for the fully disguised conditions ($p = .90$).

The results for correct rejections were similar to correct identifications. We again found a main effect of lineup type, with sequential lineups leading to significantly more correct rejections than simultaneous lineups, $F(1, 174) = 8.56, p = .004, \eta^2_p = .05$. There was also a type of disguise by degree of disguise interaction, $F(3, 174) = 3.05, p = .03, \eta^2_p = .02$. The stocking experiment led to more correct rejections than the toque and sunglasses experiment, but only for the sunglasses only vs. 2/3 covered level of disguise ($p = .05$). This is unsurprising since the 2/3 covered disguise condition does appear to be the most difficult disguise condition (see Figures 1 and 2) but does suggest that at some point witnesses do become sensitive to viewing conditions.
Finally, for target-present selections, there was a main effect of lineup type, $F(1, 174) = 8.56, p = .004, \eta^2_p = .05$ with more target-present selections from simultaneous ($M = .84, SD = .15$) than sequential lineups ($M = .79, SD = .13$), and a main effect of level of disguise, $F(3, 522) = 24.42, p < .001, \eta^2_p = .12$, with selections decreasing as targets were more disguised. There was no effect of type of disguise on selections ($p = .20$).

**General Discussion**

We expected that disguised targets would lead to fewer correct identifications and correct rejections than undisguised targets, and that there would be a linear relationship between degree of disguise and degree of accuracy, supporting the explanation offered by Brewer and colleagues (2005) that disguises reduce the information available for encoding. The results of these experiments support these intuitions and raise some interesting questions. For example, covering most of one’s face with a stocking may be as effective, or more effective, than completely covering one’s face. Why might this be? If disguises are effective because they disrupt global facial configurations then we would expect this finding. On the other hand, if disguises are effective because they mask feature information, we would expect that the full stocking disguise condition would be more effective than the 2/3 stocking disguise condition. However, as can be seen in Figure 2, the global picture of the target’s face is arguably clearer in the fully disguised condition than in the 2/3 disguise condition. The implication of these findings is that when an eyewitness views a perpetrator in disguise the likelihood of an erroneous identification depends upon not only the degree to which the perpetrator was disguised, but also which part of the face was disguised.

We found that a toque is much less disruptive to recognition than sunglasses. One reason for this difference may be that people are aware that hair can readily be changed or just vary
naturally (e.g., windblown appearance) whereas eyes are constant, resulting in greater reliance on matching eye than hair cues. Certainly much research points to the importance of eyes in face recognition and identification (e.g., Janik et al., 1978). As such, identifications of suspects when the perpetrator wore sunglasses should be viewed cautiously, and identifications of suspects when the perpetrator wore both a hat and sunglasses should be even more questionable.

Another purpose of the current experiments was to evaluate the effect of a stocking disguise relative to a toque and sunglasses disguise. It is plausible that jurors might reason that the impact of a stocking on recognition would be greater than the impact of a toque and sunglasses because it is clearly a disguise and people have less experience recognizing people wearing a stocking than a toque and/or sunglasses. Indeed as our analyses show, the stocking disguise led to fewer correct identifications compared to the toque and sunglasses disguise, overall. We conducted pairwise comparisons to clarify this finding and determined that in the most disguised condition for each type of disguise (toque and sunglasses, fully covered with stocking) there was no difference between the types of disguise but correct identifications were lower for the stocking disguise for all other conditions (only marginally so for the no disguise conditions).

This is surprising since the correct identification effect sizes for the sunglasses disguise (.39) and for the toque disguise (.20) were larger than for the stocking disguise (.16). Inspecting the means we can see that performance in the no disguise condition is responsible: participants in the toque and sunglasses experiment made an average of .87 correct identifications overall whereas participants in the stocking experiment made an average of .80 correct identifications overall. In the fully disguised condition for both experiments (i.e., toque and sunglasses or fully covered with a stocking), the mean correct identification rate was .55. Thus, the toque and
sunglasses disguise led to a larger decrease in correct identifications across degree of disguise and a higher overall average correct identification rate, potentially because participants were more willing to choose in this experiment. As discussed earlier, participant witnesses may expect a stocking to decrease their identification accuracy and so be less inclined to choose, even though its effect on accuracy is smaller. In sum, given the effect sizes, we contend that the toque and sunglasses disguise is indeed the stronger disguise.

These disguises may impact future recognition via different processes: the stocking seems to disrupt the global configuration of the target’s face with some distortion of features, whereas the toque and sunglasses obstruct the view of specific features. The results are consistent with our expectation that obscuring the eyes versus the whole face has the largest impact on accuracy. The indirect comparison of effect sizes supports the contention that the eyes are critical for face recognition as the effect of sunglasses was larger than for the toque or stocking. This further implies that identifications of a suspect after witnessing a sunglasses-disguised perpetrator should be less trusted than identifications of a suspect when the perpetrator wore a stocking or a toque, all other conditions being equal.

The results with correct rejections bear mention because they have received little attention within the disguise literature in general. We found that while the effect of disguise on correct rejections mirrored that of correct identifications (i.e., they decreased with greater disguise), the effect was not as strong. Only sunglasses elicited a significant effect. Looking at the results for confidence, this may be due to meta-cognitive effects. Perhaps the absence of a strong match to ones memory combined with the presence of a disguise interacts to raise response criterion. Future research could explore how similarity between the innocent suspect and the target influences correct rejections to try and flesh this out.
We expected to replicate the pattern of higher correct identifications and lower correct rejections with simultaneous lineups as compared to sequential lineups. This pattern was generally present even though it did not always reach significance. Moreover, the pattern was clearly present when we combined the experiments. We found no significant interactions between disguise and lineup type, suggesting that one lineup type is not more robust than the other when dealing with disguise, although the marginal interaction of type of disguise and lineup type leaves open the possibility that more correct identifications are obtained with a simultaneous than a sequential lineup if a toque and sunglasses disguise is used.

Notably, our correct identification rates were very high compared to other experiments on simultaneous and sequential lineups. The lowest average correct identification rate was still above .50. The reasons for this are twofold: (a) this was a repeated measures design in which participants were aware that they would be asked to identify the perpetrator from the video, which leads to higher correct identifications overall than if participants are unaware of the video’s purpose (Beaudry, Leach, Mansour, Bertrand, & Lindsay, 2006), and (b) the delay between exposure and lineup was almost nonexistent. Thus, the variability in viewing conditions may have been insufficient for differences due to lineup type to emerge. Further research should examine the interaction of lineup type and disguise using a between-subjects design and/or a filler task between target exposure and lineup presentation to prevent ceiling effects.

Finally, we predicted that confidence in decisions would decrease as disguise increased. This was generally true for correct identifications, though the pattern was more complex for correct rejections. The means in Table 2 show a general pattern whereby confidence decreases as the degree of disguise increases with one exception. For simultaneous lineups, participants were more confident if targets wore a toque than if they were not wearing a toque. Presumably this
occurs only for simultaneous lineups because of the increased willingness of witnesses to choose from these lineups, and the cognitive dissonance that may occur when one fails to choose. A similar effect occurred with the stocking disguise experiment such that simultaneous lineup correct rejections were made more confidently in the undisguised condition.

The results of this study supported Brewer et al.’s (2005) contention that disguise works by allowing less identifying information for encoding. If a face is fully obscured (as when a person stands behind a tree), we are not expected to be able to later recognize that face. Disguise seems to work in a similar way: disguises obscure facial information so that less information about a face is available to be encoded. Thus, there is less information to use during a recognition decision, regardless of the effect of intervening factors such as rehearsal and interference. The current research supported this by demonstrating that the more disguised a target was, the less likely decisions were to be accurate. It is important to note that this research does not rule out a role for perceived dangerousness, distraction, or encoding specificity. What this research does do is provide support for two mechanisms: availability of information for encoding and meta-cognitions. Importantly, this research also demonstrated that the effect of disguise is more complicated than just amount of information available for encoding. The lowest level of accuracy was associated with both a fully covered face and a disguise that disrupted the view of the face (2/3 stocking) and the lowest level of confidence was associated with sunglasses (the most effective disguise), which leave a large amount of the face undisguised.

Future research should explore the contribution of the various mechanisms discussed here. Eye tracking may be one way to address the extent to which witnesses are distracted by disguise and thus not able to encode the available information. Perceived dangerousness might be testable measuring galvanic skin response to events involving variously disguised individuals.
Finally, it seems clear that research on disguise should explore the meta-cognitions witnesses experience about identifying someone they witnessed in a disguise.

An important finding from this research that can be confirmed by inspecting Tables 1 and 3 is that while accurate decisions decrease with increasing disguise, choosing decreases at a much slower rate. This replicates findings by Lindsay et al. (2009) who examined accuracy and choosing as distance between the witness and the target increased during encoding. That is, witnesses do not effectively use information about quality of the encoding conditions in their decision. Social pressure to make a selection may outweigh a witness’ meta-cognitions about their ability to decide correctly. Notably, our participants completed multiple trials with variable viewing conditions which should have cued them to the importance of quality of view. Further they completed the study at a private computer terminal, in the absence of any social pressure. As such, we might expect that degree of disguise would have even less of an effect on choosing rates (but not identification rates) in the real world. Future research should consider the benefit of instructions to the witness to consider the quality of their memory or exposure to the perpetrator’s face.

A number of conclusions follow from this research. First, disguise hurts identification accuracy, though the relationship is not as simple as more coverage leads to poorer accuracy. What part of the face is covered matters. This relationship may be one means by which to compare the credibility of identifications across witnesses—for example, when multiple witnesses view a perpetrator from different locations or when witnesses see a perpetrator at different stages of a crime (e.g., disguised inside of a bank but only partially disguised or not disguised outside). Second, the impact of a sunglasses (eye covering) disguise is considerable and larger than the effect of a toque (hair covering), and when used in combination with a toque
leads to similar detriments in identification accuracy as covering a face fully or 2/3 with a stocking—even though significantly more of the face is exposed in the former condition. Third, while both accuracy and choosing decrease with quality of view, choosing decreases at a much slower rate, thereby exacerbating the negative effect of poor viewing conditions on accuracy.
References


Table 1

*Mean Identification Accuracy by Disguise and Lineup Type for Experiment 1. Standard deviations are presented in parentheses.*

<table>
<thead>
<tr>
<th></th>
<th>Target-Present Lineup</th>
<th>Target-Absent Lineup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simultaneous</td>
<td>Sequential</td>
</tr>
<tr>
<td>No disguise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.94 (.14)</td>
<td>.80 (.26)</td>
</tr>
<tr>
<td>Foil</td>
<td>.01 (.05)</td>
<td>.12 (.19)</td>
</tr>
<tr>
<td>None</td>
<td>.05 (.12)</td>
<td>.08 (.17)</td>
</tr>
<tr>
<td>Toque only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.82 (.24)</td>
<td>.74 (.24)</td>
</tr>
<tr>
<td>Foil</td>
<td>.04 (.13)</td>
<td>.09 (.19)</td>
</tr>
<tr>
<td>None</td>
<td>.13 (.20)</td>
<td>.17 (.21)</td>
</tr>
<tr>
<td>Sunglasses only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.73 (.25)</td>
<td>.64 (.31)</td>
</tr>
<tr>
<td>Foil</td>
<td>.10 (.16)</td>
<td>.14 (.19)</td>
</tr>
<tr>
<td>None</td>
<td>.16 (.24)</td>
<td>.21 (.27)</td>
</tr>
<tr>
<td>Toque &amp; sunglasses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.59 (.29)</td>
<td>.50 (.34)</td>
</tr>
<tr>
<td>Foil</td>
<td>.18 (.22)</td>
<td>.18 (.25)</td>
</tr>
<tr>
<td>None</td>
<td>.23 (.23)</td>
<td>.31 (.33)</td>
</tr>
</tbody>
</table>
Table 2

Mean percent Confidence by Disguise and Lineup Type for correct identifications and correct rejections from Experiment 1 and 2. Standard deviations are presented in parentheses.

<table>
<thead>
<tr>
<th>Disguise</th>
<th>Correct Identifications</th>
<th>Correct Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simultaneous</td>
<td>Sequential</td>
</tr>
<tr>
<td><strong>Experiment 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>85.82 (10.58)</td>
<td>52.88 (44.42)</td>
</tr>
<tr>
<td>Toque only</td>
<td>77.84 (14.98)</td>
<td>47.44 (41.00)</td>
</tr>
<tr>
<td>Sunglasses only</td>
<td>74.62 (16.26)</td>
<td>47.15 (40.20)</td>
</tr>
<tr>
<td>Toque &amp; Sunglasses</td>
<td>70.42 (17.23)</td>
<td>40.60 (38.70)</td>
</tr>
<tr>
<td><strong>Experiment 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>85.26 (12.29)</td>
<td>89.81 (11.59)</td>
</tr>
<tr>
<td>1/3 covered</td>
<td>82.12 (14.43)</td>
<td>81.54 (14.51)</td>
</tr>
<tr>
<td>2/3 covered</td>
<td>67.86 (20.41)</td>
<td>73.96 (16.75)</td>
</tr>
<tr>
<td>Fully covered</td>
<td>70.64 (15.81)</td>
<td>74.33 (17.92)</td>
</tr>
</tbody>
</table>
Table 3

*Mean Identification Accuracy by Disguise and Lineup Type for Experiment 2. Standard deviations are presented in parentheses.*

<table>
<thead>
<tr>
<th>ID</th>
<th>Target-Present Lineup</th>
<th>Target-Absent Lineup</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simultaneous</td>
<td>Sequential</td>
<td>Overall</td>
<td>Simultaneous</td>
<td>Sequential</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No disguise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.82 (.27)</td>
<td>.80 (.23)</td>
<td>.80 (.26)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Foil</td>
<td>.12 (.23)</td>
<td>.08 (.16)</td>
<td>.10 (.19)</td>
<td>.30 (.29)</td>
<td>.16 (.22)</td>
</tr>
<tr>
<td>None</td>
<td>.05 (.14)</td>
<td>.12 (.19)</td>
<td>.10 (.18)</td>
<td>.70 (.29)</td>
<td>.84 (.22)</td>
</tr>
<tr>
<td>1/3 Covered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.69 (.33)</td>
<td>.69 (.31)</td>
<td>.69 (.31)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Foil</td>
<td>.15 (.23)</td>
<td>.13 (.23)</td>
<td>.14 (.23)</td>
<td>.31 (.27)</td>
<td>.21 (.24)</td>
</tr>
<tr>
<td>None</td>
<td>.16 (.22)</td>
<td>.18 (.24)</td>
<td>.17 (.25)</td>
<td>.69 (.27)</td>
<td>.79 (.24)</td>
</tr>
<tr>
<td>2/3 Covered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.50 (.32)</td>
<td>.54 (.32)</td>
<td>.53 (.33)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Foil</td>
<td>.27 (.26)</td>
<td>.17 (.23)</td>
<td>.21 (.25)</td>
<td>.39 (.31)</td>
<td>.29 (.32)</td>
</tr>
<tr>
<td>None</td>
<td>.23 (.30)</td>
<td>.29 (.30)</td>
<td>.26 (.30)</td>
<td>.60 (.31)</td>
<td>.70 (.32)</td>
</tr>
<tr>
<td>Fully Covered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspect</td>
<td>.56 (.31)</td>
<td>.53 (.29)</td>
<td>.54 (.29)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Foil</td>
<td>.22 (.23)</td>
<td>.16 (.20)</td>
<td>.18 (.22)</td>
<td>.38 (.32)</td>
<td>.31 (.30)</td>
</tr>
<tr>
<td>None</td>
<td>.22 (.23)</td>
<td>.31 (.29)</td>
<td>.27 (.27)</td>
<td>.61 (.32)</td>
<td>.69 (.30)</td>
</tr>
</tbody>
</table>
Figure 1: Examples of stimuli used in Experiment 1. A. No disguise; B. Toque disguise; C. Sunglasses disguise; D. Toque and Sunglasses disguise.
Figure 2: Examples of stimuli used in Experiment 2. A. No disguise; B. 1/3 covered; C. 2/3 covered; D. Fully covered.