# Impact of Disguise on Identification Decision and Confidence with Simultaneous and Sequential Lineups



Jamal K. Mansour[1], Jennifer L. Beaudry[2], Michelle I. Bertrand[3], Natalie Kalmet[4], Elisabeth I. Melsom[4], Roderick. C. L. Lindsay[5]

[1] Memory Research Group; Centre for Applied Social Sciences; Psychology, Sociology, and Education; Queen Margaret University

[2] Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology

[3] Department of Criminal Justice, University of Winnipeg

[4] Independent Researchers

[5] Department of Psychology, Queen's University

## Author Note

Jamal Mansour https://orcid.org/0000-0001-7162-8493, Jennifer, L. Beaudry https://orcid.org/0000-0003-1596-6708, Michelle I. Bertrand https://orcid.org/0000-0001-5079-2358

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.

This research was supported in part by a grant from the Social Sciences and Humanities Research Council of Canada to 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.

Please address correspondence to Jamal K. Mansour, Psychology, Sociology, & Education, Queen Margaret University, Edinburgh, UK, EH21 6UU; E-mail: jmansour@qmu.ac.uk

Abstract

**Objectives:** Prior research indicates disguise negatively affects lineup identifications but the mechanisms by which disguise works have not been explored and different disguises have not been compared. We investigated how two different types of disguise, four levels of varying degrees of coverage, and lineup type influence eyewitnesses' identification decisions, accuracy, and confidence.

**Hypotheses:** We predicted that identification accuracy would decrease as the disguise covered more of a perpetrator's face. We also predicted that type of disguise—stocking mask versus sunglasses and/or toque (i.e., knitted hat)—would influence identifications, but we had conflicting predictions about which disguise would impair their performance more.

**Method:** In two experiments (*N*s = 87 and 91) we manipulated degree of coverage by two different types of disguise: a stocking mask or sunglasses and toque. Participants viewed mock-crime videos followed by simultaneous or sequential lineups.

**Results and Conclusions:** Disguise and lineup type did not interact. In support of the view that disguise prevents encoding, identification accuracy generally decreased with degree of disguise. For the stocking disguise, however, full and 2/3 coverage led to approximately the same rate of correct identifications—which suggests that disrupting encoding of specific features may be as detrimental as disrupting a whole face. Accuracy was most affected by sunglasses and we discuss the role meta-cognitions may have played. Lineup selections decreased more slowly than accuracy as coverage by disguise increased, indicating witnesses are insensitive to the effect of encoding conditions on accuracy.

*Keywords:* lineups, disguise, estimator variables, sequential lineups, simultaneous lineups

## Public Significance Statement

In two experiments we tested how differing types and degrees of disguise influence eyewitness lineup decisions and confidence. When perpetrators were disguised (versus not disguised), identifications of guilty suspects and rejections of lineups containing innocent suspects decreased. Generally, these effects were stronger when more of the perpetrator's face was covered but we also note that 1) sunglasses had a greater effect than a hat and 2) a stocking obscuring two-thirds of a perpetrator's face was as effective as one obscuring a perpetrator's entire face. Also, eyewitnesses may not fully appreciate how much a disguise influences their accuracy: choosing from lineups decreased more slowly than accuracy as degree of disguise increased. The effects of disguise on identification accuracy and confidence were similar across simultaneous and sequential lineups. Our results provide the criminal justice system, notably the courts, with a nuanced understanding of how different disguises may affect eyewitness accuracy and 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). However, this position is not one that is necessarily accepted by the criminal justice system. For example, only 45% of judges surveyed about eyewitness identification believed it would be harder to later recognize a perpetrator who was originally seen wearing a hat (50% had no opinion; Wiser & Safer, 2004). Thus, the impact of disguise on identification decisions and confidence warrants further exploration. In addition, researchers need to understand the mechanisms through which disguise affects witness decisions. Brewer, Weber, and Semmler (2005) identify two ways in which disguise may impact identification accuracy. First, viewing a disguised perpetrator compared to an undisguised one imparts less identifying information for encoding, thereby resulting in a less complete memory of the perpetrator. A second, not mutually exclusive, explanation is that disguise may affect witnesses' perceptions of the difficulty of the identification task. These meta-cognitions may increase a witness' decision criterion and subsequently decrease the likelihood of choosing 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 recognition accuracy, generally in the same form as later retrieval (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: correct identifications were highest for faces undisguised at both encoding and recognition, but worst for faces disguised at encoding while undisguised at recognition. The two remaining conditions (undisguised → disguised, and disguised → disguised) did not differ from each other. Similarly, Patterson and Baddeley (1977) found that identification accuracy (hit rate but not discriminability) varied with the match between characteristics (e.g., with or without a beard) presented at encoding and recognition, but if a face was disguised at recognition, identification accuracy declined regardless of the inclusion of disguise at encoding. In summary, encoding specificity may play a role but on its own cannot account for the influence of disguise on identification.

Finally, a fourth possibility is that disguise influences how witnesses allocate their attention. Witnesses may interpret a disguised perpetrator as more dangerous, and this perception may increase cognitive load. As a result, witnesses may have fewer processing resources available for encoding (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). Similarly, type of disguise may influence how the target's intentions are perceived (i.e., if someone looks like a "criminal" or not), and thus influence encoding strategies. Finally, a disguise may distract witnesses by drawing attention to the disguise itself and away from available facial information required for later recognition. The extant literature provides little evidence regarding which of the four mechanisms described—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). In the following experiments we explored 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.

A disguise can cover many parts of a face, but hair is particularly simple to manipulate. A perpetrator's hairstyle and/or facial hair may be easily disguised and/or changed between a crime (encoding) and the identification procedure (test), and such changes consistently decrease recognition accuracy (Cutler et al., 1987a, 1987b; Patterson & Baddeley, 1977). Moreover, obstruction of hair cues negatively impacts identification accuracy (e.g., Cutler et al., 1987a, 1987b; Narby, Cutler, & Penrod, 1996). 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 identification; however, his female target's shoulder-length hair was still visible so hair cues may not have been appreciably obscured.

Perpetrators can also easily cover their eyes. McKelvie (1976) found that wearing eyeglasses at either encoding or recognition resulted in more face recognition errors. Likewise, the addition of eyeglasses at test hinders recognition (Hockley, Hemsworth, & Consoli, 1999; Terry, 1993), possibly because the eyes are a central area of focus (Henderson, Williams, & Falk, 2005; Janik, Wellens, Goldberg, & Dell'Osso, 1978). Conversely, Patterson and Baddeley (1977) found inconsistent effects of the presence of glasses on face recognition—though they found an overall main effect. Changing the presence/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.

We wanted to directly compare the impact of disguising hair and eyes on identification accuracy, and to examine if these disguises were additive, though most disguise research separately considers the role of the eyes (e.g., McKelvie, 1976) or hair (e.g., Wright & Sladden, 2003). Some researchers claim that the eyes are the most important facial feature for recognizing a face (e.g., Henderson, et al., 2005). Research outside of the recognition literature found that participants who were instructed to form impressions of faces reported looking primarily at the eye and mouth regions (Janik et al., 1978). Given these results, we predicted that covering a target's eyes would be more detrimental to later identification than covering their hair.

Hats and sunglasses are worn by many people in everyday life, 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 threatened by the disguised individual, or be 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 may induce feelings of threat and influence the allocation of attention. A stocking disguise is conspicuous, commonly used, and is believed to be effective.

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 was poorer when targets were studied with a stocking covering their face than without a stocking. Their third experiment demonstrated that a stocking disguise

decreases recognition of targets due to the distortion of facial features, rather than the filtering out of complexion information (i.e., more global information). At this point, it is unclear whether the general flattening of facial features or the flattening of particular features reduced identification accuracy.

The current experiments further explored the effect of a stocking disguise, and investigated whether identification accuracy and selections from lineups generally decrease as disguise increases. If disguise works by decreasing the amount of information available for encoding, we should see accuracy decrease as the stocking covers more of a target's face. If meta-cognitions are completely responsible for the detrimental effect of disguise, we would expect a static negative effect when a face is disguised (to any degree) versus undisguised. However, we doubted the impact that disguise would 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 reduce their lineup selections as the to-be-identified target was presented further away, even though the accuracy of those selections decreased.

Examining various disguises (i.e., toque, sunglasses, stocking) is important because they may differentially affect recognition (Davies & Flin, 1984). That is, whether a disguise that *distorts* the global view of a face and features (stocking) may be less damaging to identification accuracy for unfamiliar faces than a disguise that *obscures* specific features (i.e., hair and eyes covered by toque and sunglasses). Figures 1D and 2D demonstrate how individual features are more discernable in the stocking disguise conditions compared to the toque and sunglasses conditions. We expected feature obstruction to be more detrimental to identification accuracy

than distortion for two reasons. First, unfamiliar faces are processed more featurally than globally (Hancock, Bruce, & Burton, 2000), 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 expect to be more accurate when they view faces with a stocking covering because they are 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 toque and sunglasses disguise.

Using the same targets and viewing scenarios across the two types of disguise (i.e., toque and sunglasses versus stocking) allowed us to compare their impact on identification accuracy. For each type, we created four levels of disguise. In Experiment 1, targets wore no disguise, a toque only (i.e., knitted hat), sunglasses only, or both at exposure. In Experiment 2, targets wore no disguise, a stocking that covered their hair and forehead, a stocking that covered their head to just below their nose, or a stocking that covered their entire head at exposure. These within-disguise manipulations enabled comparisons of the four levels of obstruction/disruption to encoding a face on identification decisions and confidence.

Both experiments also permitted a comparison of target-present and target-absent lineups across all disguise conditions and across simultaneous and sequential lineup procedures—the types commonly used in North America. 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 perpetrator 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). There is debate over which method is preferable because simultaneous lineups typically result in more correct identifications while sequential lineups typically result in more correct rejections; however, sequential lineups are more diagnostic overall (Steblay, Dysart, & Wells, 2011).

Previous disguise studies have used 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 that identification accuracy was marginally lower for simultaneous than sequential lineups, and when the target wore a hat than when he did not. The current experiments focus specifically on the issue of disguise and lineup type, and include more extensive disguise manipulations. Overall, we expected to find the standard simultaneous–sequential pattern, but we had no specific hypotheses about the interaction between disguise and lineup type.

Very little research has addressed how disguise affects target-absent lineup decisions. Cutler et al. (1987a, 1987b) and Cutler and Penrod (1988) included target-present and target-absent lineups, and although disguise significantly influenced accuracy, they did not report accuracy for target-present and -absent lineups separately. O'Rourke, Penrod, Cutler, and Stuve (1989) and Yarmey (2004) included target-absent lineups but found no effect of disguise on either target-present or target-absent lineups. One might expect disguise to lead to higher accuracy on target-absent lineups because the presence of the disguise would be a salient cue that the witness' memory may not be very detailed. As discussed previously, however, witnesses do not seem to rely on 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 predicted that disguise would similarly affect target-present and -absent lineup decisions. That is, we expected correct identifications and rejections to decrease with greater disguise because less information can be encoded—which results in less information available for the lineup decision.

The poorer the viewing conditions, the lower the confidence–accuracy relationship for lineup decisions (Bothwell, Deffenbacher, & Brigham, 1987). Some research has found that viewing a disguised versus undisguised perpetrator is akin to poor viewing conditions (e.g. O'Rourke et al., 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 examined the relationship between disguise and confidence in both experiments.

Our hypotheses are summarized here. First, we hypothesized that identification accuracy (correct identifications and correct rejections) would decrease as disguise covered more of a perpetrator's face because less information about the face would be available for encoding and/or because of witnesses' meta-cognitive beliefs about the quantity of information. Second, we had conflicting expectations about which type of disguise would have the greater impact on identification accuracy. We expected a stocking disguise to be less disruptive to accuracy because it *distorts* a perpetrator's face whereas a toque and sunglasses *obscures* a perpetrator's face. Further, because people are more commonly encountered wearing toques and sunglasses than a stocking, people's meta-cognitions about their ability to accurately identify the perpetrator may make them less willing to choose from lineups when they see a stocking-disguised target. However, the fact that a stocking allows witnesses to see features may increase their willingness

and ability to identify someone relative to a toque and sunglasses disguise. Third, we expected to replicate the standard pattern of identification performance for simultaneous and sequential lineups (higher correct identifications for simultaneous lineups and lower false identifications for sequential lineups). Finally, we explored but did not have a specific prediction about how disguise would impact confidence.

## Experiment 1

In Experiment 1 we considered the effect of toque and sunglasses disguises on lineup identification accuracy and confidence across target-present and -absent simultaneous and sequential lineups. We predicted that wearing sunglasses would be more detrimental to identification accuracy than wearing a toque, and that combining the toque and sunglasses disguises would have an additive effect. With regard to lineup type, we expected the standard pattern of more correct identifications with simultaneous lineups and more correct rejections with sequential lineups. Additionally, we wanted to know if and how this pattern would change with degree of disguise. Finally, we explored how confidence varied with degree of disguise.

### Method

**Participants.** Undergraduate students living in Ontario, Canada ($N$ = 98; 67 female, $M$ age = 19.09 years, $SD$ = 2.08) participated in this experiment for course credit or money. Most participants were of European (.78) or Asian (.15) descent.

**Design.** The design was a 2 (Lineup Type: simultaneous, sequential) x 2 (Target Presence: present, 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. Each participant saw every possible combination of disguise (4), by target sex (2) in each block yielding eight trials per

block. Across the three blocks, 12 trials were target-present and 12 were target-absent.

**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 an off-screen police officer 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, was manipulated. Approximately half of the data was collected in the 2006/07 academic year when participants viewed long and large video clips (approximately 30 seconds long, 30 cm high by 23 cm wide; when presented on a 43 cm screen), whereas the rest of the data was collected in the 2007/08 academic year when participants viewed short and small video clips (approximately 3 seconds long, and 9 cm high by 6.5 cm wide). We had no reason to expect a difference in our results as a result of data collection date. The quality of view manipulation was designed to increase the heterogeneity of the viewing conditions to ensure variability in response rates and confidence. In the interests of length, quality of view will not be discussed in detail. These interactions 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. Six-person 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. For both lineup types, z-tests indicated that the difference in the frequencies with which targets were shown in each position were not significantly different.

In order to ensure the fairness of our lineups, we recruited two additional sets of participants. The first set ($N$ = 30) provided descriptions for each target, and we randomly selected 12 of these descriptions for each target. A second independent set ($N$ = 36) engaged in a mock witness task. Each mock witness separately viewed the 24 target-present simultaneous lineups, with each lineup accompanied by one of the 12 randomly selected descriptions of that particular target. Collapsing across descriptions and lineups, Tredoux's E ranged from 2.12 to 5.23 ($M$ = 3.89, $SD$ = 0.98; Malpass, Tredoux, & McQuiston-Surrett, 2007). Given these results, we were satisfied that, in general, the 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" from the video (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. After entering their sex, age, and ethnicity, participants 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 24 trials, 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). We also wanted to ensure that participants' lineup decisions were influenced only by memory for the criminal event, and not by previous knowledge of, or exposure to, the targets. Therefore, we asked participants if they recognized someone in the lineup from real-life. 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. Police frequently ask witnesses if they know lineup members.

**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, whereas 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, whereas 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). The target-absent selection

rate is simply 1 minus the correct rejection rate.

***Confidence.*** We calculated mean confidence for correct identifications and correct rejections by disguise.

**Results**

Unlike most real-world eyewitnesses, our participants viewed 24 lineups. To determine whether learning accounted for performance on target-present and target-absent lineups, we used binomial logistic regression with trial as the predictor. For target-present lineups, target selections were coded as correct responses; all other responses were coded as incorrect. For target-absent lineups, lineup rejections were coded as correct responses and selections of any lineup member as an incorrect response. No learning effects were found ($p$s > .10).

On average, participants recognized lineup members from outside of this study from 2.56 lineups ($SD$ = 1.80; $Range$ = 1 – 9). 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 recognized a target (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 (Toque: present, absent) x 2 (Sunglasses: present, absent) x 2 (Lineup Type: simultaneous, sequential) x 2 (Quality of view of the target: good, poor) mixed-model repeated-measures analyses of variance (ANOVAs).

In the following analyses, the in-text 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 in-text mean for no toque refers to the no disguise and sunglasses only conditions. In-text means for sunglasses versus no sunglasses were calculated similarly. Table 1 presents means for the four

conditions run (i.e., no disguise, toque only, sunglasses only, toque and sunglasses). All reported confidence intervals are 95% confidence intervals.

**Correct identifications.** The correct identification rate was significantly lower when the target wore sunglasses ($M$ = .60, CI [.55, .65]) than when the target wore no sunglasses ($M$ = .82, CI [.79, .86]), $F(1, 83) = 58.52$, $p = .001$, $\eta_p^2 = .41$. Likewise, correct identifications were significantly lower when the target wore a toque ($M$ = .66, CI [.61, .70]) than when the target did not wear a toque ($M$ = .77, CI [.73, .80]), $F(1, 83) = 19.90$, $p < .001$, $\eta_p^2 = .19$. Planned comparisons showed that the correct identification rate was significantly lower when targets wore sunglasses alone versus a toque alone, $F(1, 83) = 9.59$, $p = .003$, $\eta_p^2 = .10$, and that the impact of wearing both sunglasses and toque was larger than the effect of either alone, $F(1, 83) = 26.57$, $p < .001$, $\eta_p^2 = .24$ (see Table 1 for means). Both results are consistent with our expectations. Consistent with previous research (e.g., Steblay et al., 2011), participants made more correct identifications from simultaneous ($M$ = .77, CI [.72, .82]) than from sequential lineups ($M$ = .65, CI [.61, .70]), $F(1, 83) = 12.50$, $p = .001$, $\eta_p^2 = .13$. There were no significant interactions ($p$s > .14).

**Correct rejections.** Participants made fewer correct rejections when the targets wore sunglasses ($M$ = .65, CI [.60, .70]) than when they did not ($M$ = .73, CI [.68, .78]), $F(1, 83) = 9.56$, $p = .003$, $\eta_p^2 = .10$. The effect of toque was not significant ($p = .71$); participants made just as many correct rejections when the target wore a toque ($M$ = .70, CI [.64, .75]) as when the target did not wear a toque ($M$ = .68, CI [.64, .73]). As such, we found that the effect of sunglasses was significantly greater than the effect of toque, $F(1, 83) = 4.86$, $p = .03$, $\eta_p^2 = .06$, and the combined effect of toque and sunglasses was not greater than either individually ($p$ = .55). The effect of lineup type was not significant; the correct rejection rates were similar for

sequential ($M$ = .73, CI [.67, .79]) and simultaneous lineups ($M$ = .65, CI [.59, .71]), $F(1, 83)$ = 3.47, $p$ = .07, $\eta_p^2$ = .04. Again, there were no significant interactions ($p$s > .19). 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 .77 (CI [.72, .82]) selections (correct identifications and foil selections) when targets wore sunglasses, compared with .89 (CI [.86, .92]) when they did not, $F(1, 83) = 20.98$, $p < .001$, $\eta_p^2$ = .20. For target-present toque trials, participants made an average of .79 (CI [.74, .83]) selections when the target had a toque compared to .87 (CI [.83, .90]) when the target had no toque, $F(1, 83) = 16.38$, $p < .001$, $\eta_p^2$ = .16. The main effect of lineup type on target-present selections was not-significant, $F(1, 83) = 3.77$, $p$ = .056, $\eta_p^2$ = .04; the selection rate for simultaneous lineups ($M$ = .86, CI [.81, .90]) was not significantly different from sequential lineups ($M$ = .80, CI [.75, .84]). There were no significant interactions ($p$s > .18). 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 ($M$ = 70.57%, CI [67.52, 73.62]) than when the target did not wear sunglasses ($M$ = 80.93%, CI [78.12, 83.73]), $F(1, 67) = 62.55$, $p < .001$, $\eta_p^2$ = .48. Likewise, confidence was lower when the target wore a toque ($M$ = 72.25%, CI [69.01, 75.49]) than when the target did not wear a toque ($M$ = 79.25%, CI [76.51, 81.99]), $F(1, 67) = 22.86$, $p < .001$, $\eta_p^2$ = .25. Overall, confidence did not significantly differ for simultaneous ($M$ = 76.36%, CI [73.01, 79.72]) versus sequential lineups ($M$ = 75.13%, CI [71.02, 79.16]), $F(1, 67) = 0.22$, $p$ = .64, $\eta_p^2$ = .003. The main effects of sunglasses and toque were qualified by a significant interaction, $F(1, 67) = 4.60$, $p$ = .036, $\eta_p^2$ = .06. There was a significant difference in confidence ratings between toque and no toque when the actor wore no sunglasses (9.79% mean difference, $p < .001$) but not

when the actor wore sunglasses (4.22% mean difference, $p = .06$). Sunglasses also interacted with quality of view , $F(1, 67) = 7.30$, $p = .01$, $\eta_p^2 = .10$, whereby the effect of sunglasses was weaker in the good ($\eta_p^2 = .22$) than poor view ($\eta_p^2 = .40$) condition, although it was significant in both ($p$s < .001). There were no other significant interactions ($p$s > .30).

**Confidence in correct rejections.** Confidence in correct rejections was lower when the target wore sunglasses ($M = 68.20\%$, CI [64.87, 71.53]) than when the target did not wear sunglasses ($M = 75.28\%$, CI [71.86, 78.69]), $F(1, 70) = 20.09$, $p < .001$, $\eta_p^2 = .22$. There was no effect of toque ($p = .73$). There was no main effect of lineup type on confidence in correct rejections, $F(1, 70) = 0.02$, $p = .90$, $\eta_p^2 < .001$ (also see Table 2). No interactions reached significance ($p$s > .053).

**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. Overall, the effects of disguise on target-absent lineups were similar to target-present lineups, such that accuracy decreased when a target was disguised. The target-absent effects, however, were limited to the sunglasses disguise, and were less pronounced than the target-present lineups.

As expected, simultaneous lineups resulted in more correct identifications than sequential

lineups; however, sequential lineups produced similar correct rejections to simultaneous lineups. Thus, our results partially replicate typical findings with these lineups types (Steblay et al., 2011).

The findings for confidence were somewhat 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. Indeed, for correct rejections, confidence was affected by sunglasses but not a toque; people may expect their performance to be affected when eyews are obscured, but not hair. However, we found an interaction of sunglasses and toque which perhaps suggests a logarithmic or threshold effect on confidence in correct identifications: a toque led to lower confidence only compared to no disguise at all but the addition of a toque did not reduce confidence when the target was already wearing sunglasses.

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

## Experiment 2

Experiment 2 assessed 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 change the nature of the features viewed during encoding versus recognition. Different from the classic lineup paradigm however, Davies and Flin presented participants with a series of four faces and later asked them to choose these four targets from a target-present array of 16 faces. We utilized both target-present and target-

absent simultaneous and sequential lineups to replicate and extend their work on the stocking disguise. As well, we examined the effectiveness of a partial disguise. That is, compared to covering the face entirely, how is identification accuracy affected when a stocking partially covers the target's face, thereby allowing encoding of some features?

## Method

**Participants.** Undergraduate university students in Ontario, Canada ($N$ = 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: present, absent) x 4 (Disguise: none, 1/3, 2/3, full [i.e., stocking covering all of the head]) mixed design, with lineup type as the between-subjects factor, and disguise and target presence as the within-subjects factors. The nature of the disguise manipulation is described below.

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

***Videos.*** Videos of the same 24 targets from Experiment 1 served as the stimuli in this experiment. Instead of the toque and sunglasses disguise, targets were filmed with 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. As in Experiment 1, we manipulated the on-screen video sizes and video durations; this quality of view variable did

not produce significant interactions with any of the measures.

***Lineups.*** As in Experiment 1, target location varied such that targets appeared in each position in the simultaneous and sequential lineups approximately equally.

**Results**

Binomial logistic regression again confirmed that no learning effects occurred (*p*s > .10). Participants indicated prior knowledge of one or more lineup members on an average of 2.60 trials (*SD* = 1.33; *Range*: 1–7). Participants who indicated recognition of a target from outside the task (*N* = 11) were dropped from analysis. We tested 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 for the remaining 91 participants with 4 (Disguise: none, 1/3, 2/3, fully covered) x 2 (Lineup Type: simultaneous, sequential) x 2 (Quality of view of the target: good, poor) 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, $F(3, 261) = 20.87$, $p < .001$, $\eta_p^2 = .19$ (see Table 3). The expected main effect of lineup type was not found ($p = .13$); correct identifications were approximately equal when participants saw simultaneous lineups ($M = .68$, CI [.63, .74]) compared to sequential lineups ($M = .63$, CI [.58, .68]). There were no significant interactions (*p*s > .07).

**Correct rejections.** There was a main effect of disguise on correct rejections, $F(3, 261) = 6.34$, $p < .001$, $\eta_p^2 = .07$ (see Table 3). When the actor wore no disguise or the stocking covered only the top 1/3 of their face, participants made correct rejections at a similar rate ($p = .65$), which was significantly higher than when the actor wore a stocking covering 2/3 or all of their

face (*p*s < .01), which did not differ from each other (*p* = .95). Also, the expected main effect of lineup type was present: participants who viewed sequential lineups (*M* = .75, CI [.69, .80]) made significantly more correct rejections than participants who viewed simultaneous lineups (*M* = .66, CI [.60, .73]), $F(1, 87) = 4.05$, $p = .047$, $\eta_p^2 = .05$. None of the interactions were significant (*p*s > .46).

**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, $F(3, 261) = 12.42$, $p < .001$, $\eta_p^2 = .12$ (see Table 3). The main effect of lineup type was not significant (*p* = .47); simultaneous lineups led to similar selections (*M* = .80, CI [.74, .85]) as sequential lineups (*M* = .77, CI [.72, .82]). Again, no interactions were significant (*p*s > .053).

**Confidence in correct identifications**. Confidence in correct identifications decreased as degree of disguise increased, such that participants were most confident when targets wore no disguise and least confident when targets had their face 2/3 or fully covered, which did not differ, $F(3, 195) = 27.01$, $p < .001$, $\eta_p^2 = .29$ (see Table 2). There was no main effect of lineup type (*p* = .38) and no interactions (*p*s > .42).

**Confidence in correct rejections.** There was a main effect of disguise on confidence in correct rejections, $F(3, 210) = 7.98$, $p < .001$, $\eta_p^2 = .10$. The two least disguised conditions (no disguise and 1/3 disguise) led to significantly higher confidence than the two most disguised conditions (2/3 disguise and fully disguised; *p*s $\leq$ .006). There were no other significant differences (*p*s > .58; see Table 2). There was no significant main effect of lineup type (*p* = .70), but there was a disguise by lineup type interaction, $F(3, 210) = 3.85$, $p = .01$, $\eta_p^2 = .05$. Pairwise comparisons indicated that participants were significantly more confident in correct rejections

for sequential than simultaneous lineups for the 2/3 covered condition only ($p$ = .007; all other $p$s > .41; see Table 2).

**Discussion**

As in Experiment 1, and consistent with Davies and Flin (1984), disguise led to a significant reduction in correct identifications and target-present selections, but also 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 finding is consistent with past research which indicates that the eyes are most important facial feature for decisions related to impression formation, recognition, and identification (Janik et al., 1978; McKelvie, 1976; Henderson et al., 2005). Alternatively, the salient line across the target's face in the 2/3 condition distorted the global facial appearance, which may have led to similar correct identifications, target-present selection rates , and correct rejections between the 2/3 and fully covered disguise conditions. Certainly the 2/3 and fully covered conditions appear to have had the most effect on confidence across the different levels of disguise and target-presence. The current results provide further support for Davies and Flin's (1984) suggestion that a stocking disguise decreases 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, although in the 2/3 covered condition we surely disrupted features as well as the global appearance. The full disguise condition produced similar performance rates as the 2/3 condition; apparently, as long as the hair, eyes, and nose are disrupted, disruption of the mouth and chin is unnecessary. Further research is needed to understand the role of each feature, perhaps by tracking participants' gaze when examining faces with various types of disguise.

In Experiment 1, sunglasses (but not toque) had a significant impact on correct rejections,

with sunglasses (cf. no sunglasses) reducing correct rejections. In the current experiment, the stocking disguise also had a significant effect on correct rejections. Participants were very willing to make an identification in general—the overall rate of selections was .54—and it seems that witnesses may be unwilling to reject lineups unless the most important information for identification, the eyes, are obviously disguised. 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.

The influence of disguise on confidence in correct rejections was again very interesting. In Experiment 1, confidence was reduced when targets wore sunglasses but not when they wore a toque, and disguise did not interact with lineup type. . In Experiment 2, however, confidence in correct rejections was higher for sequential than simultaneous lineups if the target had 2/3 of their face covered. Together with the findings that confidence in correct identifications were lower for targets who wore sunglasses or toques, these results imply that confidence may be somewhat sensitive to viewing conditions. They further suggest that in some circumstances, participants may be more sensitive to poor viewing conditions when shown sequential than simultaneous lineups.

The discrepancy in results across the two experiments may reflect the nature of the two disguises. As discussed earlier, stockings (Experiment 2) are somewhat more conspicuous than a

toque and sunglasses (Experiment 1). As a result, participants exposed to the stocking disguise may have expected that it would negatively impact their ability to accurately select someone from the lineup. Participants exposed to the less conspicuous toque and sunglasses disguise may not have had the same expectation.

## 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 statistically compare the effects of the two disguises. 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.

For correct identifications, there was a significant main effect of lineup type. Correct identifications were higher for participants who viewed simultaneous ($M$ = .73, CI [.70, .77]) compared to sequential lineups ($M$ = .65, CI [.62, .68]), $F(1, 174) = 11.60$, $p = .001$, $\eta_p^2 = .06$. The main effect of disguise was significant, $F(3, 522) = 46.64$, $p < .001$, $\eta_p^2 = .21$, 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, CI [.68, .75]; Experiment 1) than in the stocking experiment ($M$ = .67, CI [.63, .70]; Experiment 2), $F(1, 174) = 4.62$, $p = .033$, $\eta_p^2 = .03$. We will further discuss the impact of disguise type in the general discussion.

These main effects were qualified by an interaction. There was a significant interaction of type of disguise and degree of disguise for correct identifications, $F(3, 522) = 2.63$, $p = .050$, $\eta_p^2 = .02$. Which disguise was worn did not affect correct identifications in the no disguise condition ($p = .20$). More important to note, the toque-and-sunglasses disguise led to significantly more correct identifications than the stocking in one of the two intermediate disguise conditions (sunglasses only vs. 2/3 covered condition, $p = .009$). There was no difference in correct identifications for the toque only vs. 1/3 covered condition ($p = .085$) or for the fully disguised conditions ($p = .50$).

The results for correct rejections were similar to correct identifications. For correct rejections, we again found a main effect of lineup type, with sequential lineups leading to more correct rejections than simultaneous lineups, $F(1, 174) = 8.24$, $p = .005$, $\eta_p^2 = .04$. The main effect of degree of disguise was significant, $F(3, 522) = 9.86$, $p < .001$, $\eta_p^2 = .05$. The two lesser disguise conditions (i.e. no disguise and 1/3 coverage stocking/toque only) did not differ significantly ($p = .66$). Likewise, the two greater disguise conditions (sunglasses only/ 2/3 coverage stocking and full disguise) did not differ significantly ($p = .74$). All other pairwise comparisons were significant ($ps \leq .001$).

Finally, for target-present selections, there was no significant effect of lineup type, $F(1, 174) = 3.21$, $p = .075$, $\eta_p^2 = .02$, with similar target-present selections from simultaneous ($M = .83$, CI [.80, .86]) and sequential lineups ($M = .79$, CI [.76, .82]). There was a significant main effect of level of disguise, $F(3, 522) = 24.33$, $p < .001$, $\eta_p^2 = .12$, with selections decreasing as targets were more disguised. There was also no significant effect of type of disguise on selections, $F(1, 174) = 3.30$, $p = .071$, $\eta_p^2 = .02$; participants made similar selections if they had seen the toque and sunglasses disguises ($M = .83$, CI [.80, .86]) and the stocking disguises ($M =$

.79, CI [.76, .82]).

## General Discussion

Brewer et al. (2005) suggested that disguises obscure facial information so that less of this information is available to be encoded. As a result, witnesses have less information to use for recognition, regardless of intervening factors such as rehearsal and interference. Consistent with this hypothesis, we expected fewer correct identifications and correct rejections with disguised than undisguised targets and that there would be a linear relationship between degree of disguise and degree of accuracy. Our data support these hypotheses: the more disguised a target was, the less likely participants were to make an accurate lineup decision. However, the data also raise an interesting issue: covering most of a face with a stocking was as effective as completely covering it. This result is consistent with the explanation that disguise is effective because it disrupts global facial configurations. If disguises were effective because they mask feature information (Davies & Flin, 1984), the full stocking disguise condition should have been more effective than the 2/3 stocking condition. However, the global picture of the target's face is arguably clearer in the fully disguised condition than in the 2/3 disguise condition (see Figure 2). The implication is that, when an eyewitness views a perpetrator in disguise, the likelihood of an erroneous identification depends not only on the degree to which the perpetrator was disguised but also on which part of the face was disguised.

Indeed, we found that a toque is less disruptive to recognition than sunglasses. One reason 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., Henderson, et al., 2005; Janik et al., 1978). Identifications of suspects when

the perpetrator wore sunglasses should be viewed cautiously, and identifications of suspects when the perpetrator wore a hat and sunglasses should be even more questionable.

One purpose of this research was to evaluate disguise with a stocking relative to a toque and sunglasses. Jurors may reason that a stocking has a greater impact on recognition than a toque and sunglasses because a stocking is clearly a disguise and people have less experience recognizing people wearing a stocking than a toque and/or sunglasses. Moreover, when witnesses encounter perpetrators in a stocking disguise, they may be more inclined to try and encode their face, but may be less confident in their ability to do so. Indeed, overall, the stocking disguise resulted in fewer correct identifications than the toque and sunglasses disguise. In the most disguised conditions (toque and sunglasses, fully covered with stocking), the second least disguised condition (toque only, 1/3 covered), and the no disguise condition, there was no difference between the types of disguise, but correct identifications were significantly lower for the 2/3 covered than the sunglasses only condition. Thus, a conspicuous disguise (stocking) leads to poorer recognition than an inconspicuous one (toque or sunglasses), but conspicuousness is irrelevant when the disguise fully obscures the face.

The findings just discussed are surprising since the correct identification effect size~~s~~ for the sunglasses disguise (.41) was larger than for the stocking disguise (.19) and the toque only disguise (.19). Inspecting the means we can see that performance was better in Experiment 1 for all conditions except the fully disguised condition, where performance was nearly identical (.54 in Experiment 1, .56 in Experiment 2). Thus, the toque-and-sunglasses disguise produced a higher, overall average, correct identification rate and led to a larger decrease in correct identifications across degree of disguise; potentially, participants were more willing to choose in this experiment overall. As discussed earlier, witnesses may expect a stocking to decrease their

identification accuracy and so be less inclined to choose. In sum, we contend that the toque-and sunglasses disguise is the stronger disguise because its negative effects on recognition accuracy are similar to the stocking disguise, but it may have less impact on a witness' belief in their ability to choose. However, the similarity of target-present lineup selections across all disguise conditions is at odds with this explanation. Clearly, the role of metacognitions and disguises should be studied in greater depth.

A possible reason for the different results between experiments is that our disguises 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 a larger impact on accuracy. The effect size comparison 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.

Correct rejections bear mention because they have received little attention within the disguise literature in general. The effect of disguise on correct rejections mirrored that of correct identifications (i.e., decreased with greater disguise), the effect was significant for sunglasses and the stocking disguises, but not the toque disguise. Again, meta-cognitions may be at play, as supported by changing confidence levels. Perhaps the absence of a strong match to one's memory combined with the presence of a disguise interacts to raise response criterion. Future

research could explore how similarity between an innocent suspect and the target influences correct rejections to flesh this out.

We expected higher correct identifications and lower correct rejections with simultaneous compared to sequential lineups. This pattern was present though it did not always reach significance in the individual experiments. We found no significant interactions between disguise and lineup type (for lineup decisions), suggesting that one lineup type is no more robust than the other when dealing with disguise. A caveat is that ceiling effects may have prevented an interaction of disguise and lineup type from emerging. We suggest a between-subjects design and/or a filler task between target exposure and lineup presentation to further examine this relationship.

Finally, we were interested in how confidence varied with disguise. Confidence generally decreased with disguise for correct identifications but the pattern was more complex for correct rejections. Confidence in correct rejections also generally decreased as disguise increased but not for the toque only disguise; also, confidence in correct rejections for sequential lineups was higher than for simultaneous lineups for the 2/3 stocking disguise condition.

Confidence is a critically important variable when it comes to eyewitness testimony as jurors rely heavily on this variable when determining whether to believe a witness (Cutler et al., 1990). Moreover, police officers and prosecutors often base decisions about whether to pursue a case on the confidence of the eyewitness. That confidence generally decreases with disguise indicates that witnesses are sensitive to some degree of their ability to make accurate identifications, and this sensitivity is in a form understood by triers of fact (i.e., confidence). However, the data for were not completely straightforward for correct rejections. Confidence in a lineup rejection provides evidence that a suspect is innocent and the more confidently a witness

does this, the less likely it may be that a suspect will be further investigated. However, our data suggest that meta-cognitions may interact with confidence judgements in complex ways. Only further research can tease these effects apart.

This research has notable limitations. First, this experimental design did not allow us to differentiate between the four explanations for the effects of disguise; rather we provided evidence that one of those explanations is very plausible (availability of information to encode) and suggested reasons why another likely plays a role (meta-cognitions). Future research should explore these explanations more directly. Second, we compared the effects of two particular types of disguise presented under similar conditions. These findings may not generalize to other disguises or when a disguised perpetrator is witnessed under different circumstances (e.g., witness is directly involved or has a poor view of the perpetrator, disguised or not). Third, our design had low ecological validity. We used a repeated trials design in order to control for individual differences in decision criterion and so that we could test a range of disguises with appropriate power. As such participants were aware they would have to make an identification which leads to higher accuracy overall (Beaudry, Leach, Mansour, Bertrand, & Lindsay, 2006). Furthermore, using a repeated measures design restricts the analysis of our data in consequential ways. Participants who did not provide data in all possible categories were not included in the ANOVAs conducted (e.g., someone who did not make a correct rejection in all four stocking disguise conditions). While a repeated measures design controls for individual differences, sometimes the means used for analysis are discrepant from those calculated individually. Fourth, a real-world witnessing episode is unlikely to involve viewing multiple successive crimes or completing a lineup identification immediately after witnessing the target event. Both may have led to a higher willingness to choose from lineups. Another concern of repeated trials is that

participants may have been able to discern the number of images presented in the sequential lineup, which may have increased the pressure to choose. We did not find any systematic effects on response rates across trials, however. Fifth, performance was quite good, even under our poorest of conditions, with correct identifications ranging from .48 to .94, and correct rejections ranging from .59 to .83. The pattern of results may vary considerably when performance is poorer—an issue that should be explored with future research. Sixth, this research does not examine other potentially important factors that could be related to disguise generally or to a particular disguise such as perceived dangerousness, distraction, or encoding specificity. Perceived dangerousness might be testable measuring galvanic skin response to events involving variously disguised individuals. Eye tracking could be used to study how witnesses attend to (and thus, presumably encode) disguised perpetrators.

An important finding from this research is that while accuracy decreases 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 a witness and a target increased during encoding. Witnesses did not effectively use information about quality of the encoding conditions in their decision. Social pressure to make a selection may outweigh witnesses' meta-cognitions about their ability to decide correctly. Indeed, 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. Hence, 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. In the future, researchers should consider the benefit of instructions to witnesses directing them to consider the quality of their memory or their exposure to the perpetrator's face.

A number of conclusions follow from this research. First, disguise hurts identification accuracy but the relationship is not as simple as more coverage leading to poorer accuracy. The specific features, and how they are covered, matters. The lowest levels of identification accuracy were associated with a fully covered face and a disguise that disrupted the view of the face (2/3 stocking). As a result, one may be able 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 a bank but only partially disguised or not at all disguised outside). Second, the impact of a sunglasses (eye covering) disguise is considerable and larger than the effect of a toque (hair covering). In combination, sunglasses and a toque lead to detriments in identification accuracy similar to those yielded either by covering a face fully or by 2/3 with a stocking—even though the toque and sunglasses are less conspicuous than the stocking and leave more of the face exposed. Third, while both accuracy and choosing decrease as quality of view declines, choosing decreases at a much slower rate which exacerbates the negative effect of poor viewing conditions on accuracy. Finally, confidence generally decreases when targets are disguised but again, the story is not that simple. Whether a simultaneous or sequential lineup is used influences the magnitude of confidence ratings for correct rejections. Taken together, our results suggest that identifications of disguised perpetrators should be treated cautiously and that future research is needed to more fully understand how and when disguise works.

References

Beaudry, J. L., Leach, A.-M., Mansour, J. K., Bertrand, M., & Lindsay, R. C. L. (2006, March). *The element of surprise: The impact of participants' knowledge of a subsequent lineup task.* Presentation to the American Psychology-Law Society, St. Petersburg, FLA, USA.

Bond, N. W., & McConkey, K. M. (1995). Information retrieval: Reconstructing faces. In N. Brewer, & C. Willson (Eds.), *Psychology and Policing* (pp. 101 - 117). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bothwell, R. K., Deffenbacher, K. A., & Brigham, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology, 72,* 691 – 695. doi: 10.1037/0021-9010.72.4.691

Brewer, N., Weber, N. & Semmler, C. (2005). Eyewitness Identification. In N. Brewer, & K. D. Williams (Eds.), *Psychology and Law: An Empirical Perspective* (pp. 177 - 221). New York, NY: The Guilford Press.

Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73,* 281 – 290. doi: 10.1037/0021-9010.73.2.281

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14,* 185 – 191. doi: 10.1007/BF01062972

Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987a). Improving the reliability of eyewitness identifications: Putting context into context. *Journal of Applied Psychology, 72,* 629-637. doi: 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. doi:

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.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28,* 687 – 706. doi: 10.1007/s10979-004-0565-x

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47,* 857-858.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Science, 4,* 330 – 337. doi:10.1016/S1364-6613(00)01519-9

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98-106. doi: I: 10.3758/BF03195300

Hockley, W. W., Hemsworth, D. H., & Consoli, A. (1999). Shades of the mirror effect: Recognition of faces with and without sunglasses. *Memory & Cognition, 27*(1), 128 – 138. doi: 10.3758/BF03201219

Lindsay, R. C. L., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. R. (2009). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behavior, 32,* 526 – 535. doi: 10.1007/s10979-008-9128-x

Lindsay, R. C. L. & Turtle, J. (1999, October). Best practices for suspect identification: Blind sequential lineup procedures. Criminal Investigative Best Practices Symposium, Ottawa, Ontario.

Lindsay, R.C.L. & Wells, G.L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentations. *Journal of Applied Psychology, 70,*

556-564. doi: 10.1037/0021-9010.70.3.556

Malpass, R. S., Tredoux, C., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of Eyewitness Psychology: Memory for People* (pp. 155 - 178). Mahwah, NJ: Lawrence Erlbaum Associates.

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *The American Journal of Psychology, 89,* 311-323. doi: 10.2307/1421414

Narby, D. J., Cutler, B. R., & Penrod, S. (1996). The effects of witness, target, and situational factors on eyewitness identifications. In S. L. Sporer, R. S. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification* (pp. 23–52). Mahwah, NJ: Lawrence Erlbaum Associates.

O'Rourke, T. E., Penrod, S. D., Cutler, B. L. & Stuve, S. E. (1989). The external validity of eyewitness identification research: Generalizing across subject populations. *Law and Human Behavior, 13,* 385 – 395. doi: 10.1007/BF01056410

Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 406 - 417. doi: 10.1037/0278-7393.3.4.406

Sadr, J., Jarudi, I., & Sinha, P. (2003). The role of eyebrows in face recognition. *Perception, 32,* 285-293. doi: 10.1068/p5027

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

Steblay, N., Dysart, J., & 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. doi: 10.1037/a0021650

Terry, R. L. (1993). How wearing eyeglasses affects facial recognition. *Current Psychology, 12*, 151-162. doi: 10.1007/BF02686820

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373. doi: 10.1037/h0020071

van Koppen, P. J., & Lochun, S. K. (1997). Portraying perpetrators: The validity of offender descriptions by eyewitnesses. *Law and Human Behavior, 21,* 661 – 685. doi: 10.1023/A:1024812831576

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*(12), 1546-1557. doi: 10.1037/0022-3514.36.12.1546

Wiser, R. A., & Safter, M. A. (2004). What US judges know and believe about eyewitness testimony. *Applied Cognitive Psychology, 18*, 427-443. doi: 10.1002/acp.993

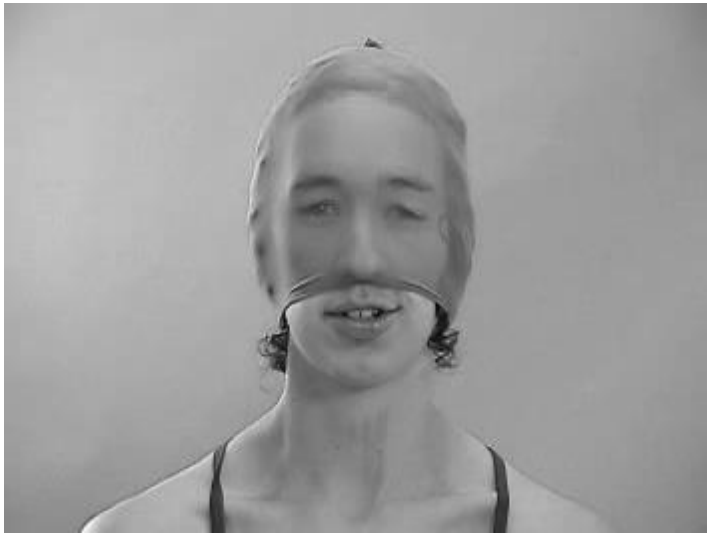
Wright, D.B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica, 114*, 101-114. doi: 10.1016/S0001-6918(03)00052-0

Yarmey, A.D. (2004). Eyewitness recall and photo identification: A field experiment. *Psychology, Crime & Law, 10*, 53-68. doi: 10.1080/1068316021000058379


*Figure 1*: Examples of stimuli from Experiment 1. A. No disguise; B. Toque disguise; C. Sunglasses disguise; D. Toque and Sunglasses disguise.

A

B



C

D

*Figure 2*: Examples of stimuli used in Experiment 2. A. No disguise; B. 1/3 covered; C. 2/3 covered; D. Fully covered.

Table 1

*Mean Identification Accuracy by Disguise and Lineup Type for Experiment 1 (N = 87).*

| ID | Target-Present Lineups | | | | | | Target-Absent Lineups | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Simultaneous | | Sequential | | Overall | | Simultaneous | | Sequential | | Overall | |
| | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI |
| No disguise | | | | | | | | | | | | |
| Suspec | .94 (.15) | [.90, .98] | .81 (.26) | [.73, .89] | .88 (.22) | [.83, .92] | NA | | NA | | NA | |
| Foil | .01 (.05) | [0, .02] | .10 (.17) | [.05, .16] | .05 (.13) | [.02, .08] | .29 (.24) | [.22, .36] | .23 (.27) | [.15, .31] | .26 (.26) | [.21, .31] |
| None | .05 (.12) | [.02, .09] | .09 (.18) | [.03, .14] | .07 (.15) | [.04, .10] | .71 (.24) | [.64, .78] | .77 (.27) | [.69, .85] | .74 (.26) | [.69, .79] |
| Toque only | | | | | | | | | | | | |
| Suspec | .82 (.24) | [.75, .89] | .75 (.24) | [.67, .82] | .78 (.24) | [.73, .84] | NA | | NA | | NA | |
| Foil | .04 (.13) | [.00, .08] | .08 (.19) | [.02, .14] | .06 (.16) | [.03, .10] | .32 (.32) | [.22, .41] | .21 (.28) | [.12, .29] | .26 (.30) | [.20, .33] |
| None | .13 (.21) | [.07, .19] | .17 (.21) | [.11, .24] | .15 (.21) | [.11, .20] | .68 (.32) | [.59, .78] | .79 (.28) | [.71, .88] | .74 (.30) | [.67, .80] |
| Sunglasses only | | | | | | | | | | | | |
| Suspec | .73 (.25) | [.66, .81] | .62 (.32) | [.52, .72] | .68 (.29) | [.62, .74] | NA | | NA | | NA | |
| Foil | .10 (.16) | [.06, .15] | .17 (.21) | [.10, .23] | .13 (.19) | [.09, .17] | .41 (.30) | [.32, .50] | .31 (.29) | [.22, .40] | .36 (.30) | [.30, .42] |
| None | .16 (.25) | [.09, .24] | .21 (.29) | [.12, .30] | .19 (.27) | [.13, .24] | .59 (.30) | [.50, .68] | .69 (.29) | [.60, .78] | .64 (.30) | [.58, .70] |
| Toque & sunglasses | | | | | | | | | | | | |
| Suspec | .59 (.30) | [.50, .68] | .48 (.34) | [.38, .59] | .54 (.32) | [.47, .61] | NA | | NA | | NA | |
| Foil | .18 (.23) | [.11, .24] | .20 (.26) | [.12, .28] | .19 (.24) | [.14, .24] | .36 (.31) | [.26, .45] | .29 (.31) | [.20, .39] | .32 (.31) | [.26, .39] |
| None | .23 (.23) | [.16, .30] | .32 (.32) | [.22, .41] | .27 (.28) | [.21, .33] | .64 (.31) | [.55, .74] | .71 (.31) | [.61, .80] | .67 (.31) | [.61, .74] |

*Note:* Means are for the conditions run (i.e., no disguise, toque only, sunglasses only, toque and sunglasses), rather than the estimated marginal means from conducted analyses of variance 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. CI = confidence interval; NA = not applicable.

Table 2

*Mean Percent Confidence by Disguise and Lineup Type for Correct Identifications and Correct Rejections From Experiment 1 and 2.*

| | Correct Identifications | | | | | | Correct Rejections | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Disguise | Simultaneous | | Sequential | | Overall | | Simultaneous | | Sequential | | Overall | |
| | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI |
| | | | | | | Experiment 1 | | | | | | |
| None | 85.82 (10.58) | [82.73, 88.91] | 87.42 (11.87) | [83.83, 91.01] | 86.58 (11.17) | [84.24, 88.93] | 75.49 (16.77) | [70.59, 80.39] | 77.48 (21.25) | [71.05, 83.90] | 76.46 (19.00) | [72.47, 80.45] |
| Toque only | 77.84 (14.98) | [73.46, 82.22] | 77.24 (14.35) | [72.90, 81.58] | 77.55 (14.59) | [74.48, 80.61] | 77.41 (15.68) | [72.83, 81.99] | 70.76 (21.17) | [64.35, 77.16] | 74.08 (18.81) | [70.13, 78.04] |
| Sunglasses only | 74.62 (16.26) | [69.87, 79.37] | 74.28 (14.02) | [70.04, 78.52] | 74.46 (15.17) | [71.28, 77.65] | 68.26 (17.45) | [63.16, 73.36] | 70.26 (21.02) | [63.90, 76.61] | 69.26 (19.22) | [65.22, 73.30] |
| Toque & Sunglasses | 70.42 (17.23) | [65.38, 75.45] | 69.05 (15.84) | [64.26, 73.84] | 69.82 (16.53) | [66.34, 73.29] | 69.17 (14.98) | [64.79, 73.54] | 67.58 (19.00) | [61.84, 73.33] | 68.39 (16.96) | [64.83, 71.96] |
| | | | | | | Experiment 2 | | | | | | |
| None | 84.35 (12.89) | [80.36, 88.34] | 89.81 (11.59) | [86.63, 92.99] | 87.44 (12.40) | [84.90, 89.99] | 77.54 (17.12) | [72.24, 82.84] | 74.26 (21.89) | [68.25, 80.26] | 75.67 (19.93) | [71.58, 79.77] |
| 1/3 covered | 81.29 (14.01) | [76.95, 85.64] | 81.54 (14.51) | [77.56, 85.52] | 81.43 (14.21) | [78.51, 84.35] | 75.10 (17.58) | [69.65, 80.55] | 73.26 (23.51) | [66.80, 79.71] | 74.06 (21.05) | [69.73, 78.38] |
| 2/3 covered | 68.57 (20.20) | [62.31, 74.83] | 73.96 (16.75) | [69.36, 78.56] | 71.61 (18.41) | 67.83, 75.40] | 60.46 (18.42) | [54.74, 66.16] | 71.90 (20.10) | [66.38, 77.41] | 66.86 (20.09) | [62.73, 70.98] |
| Fully covered | 70.99 (16.09) | [66.00, 75.98] | 74.33 (17.92) | [69.42, 79.25] | 72.82 (17.10) | [69.31, 76.33] | 64.43 (22.10) | [57.58, 71.28] | 62.56 (22.80) | [56.30, 68.82] | 63.36 (22.39) | [58.76, 67.96] |

*Note*: Means are for the conditions run, rather than the estimated marginal means from conducted analyses of variance. CI = confidence interval

Table 3

*Mean Identification Accuracy by Disguise and Lineup Type for Experiment 2 (N = 92).*

| ID | Target-Present Lineups | | | | | | Target-Absent Lineups | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Simultaneous | | Sequential | | Overall | | Simultaneous | | Sequential | | Overall | |
| | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI | *M* (*SD*) | 95% CI |
| | | | | | | No disguise | | | | | | |
| Suspect | .87 (.24) | [.79, .94] | .80 (.23) | [.73. .86] | .83 (.24) | [.78, .88] | NA | | NA | | NA | |
| Foil | .04 (.11) | [.01, .08] | .08 (.16) | [.04, .12] | .06 (.14) | [.03, .09] | .27 (.28) | [.18, .36] | .17 (.24) | [.10, .24] | .21 (.26) | [.16, .27] |
| None | .09 (.21) | [.02, .16] | .12 (.19) | [.07, .18] | .11 (.20) | [.07, .15] | .73 (.28) | [.64, .81] | .83 (.24) | [.76, .90] | .78 (.26) | [.73, .84] |
| | | | | | | 1/3 Covered | | | | | | |
| Suspect | .73 (.28) | [.65, .82] | .69 (.31) | [.61, .78] | .71 (.30) | [.65, .77] | NA | | NA | | NA | |
| Foil | .08 (.18) | [.02, .14] | .13 (.23) | [.07, .19] | .11 (.21) | [.06, .15] | .27 (.25) | [.19, .34] | .21 (.23) | [.14, .27] | .23 (.24) | [.18, .28] |
| None | .19 (.20) | [.12, .25] | .18 (.28) | [.10, .25] | .18 (.25) | [.13, .23] | .73 (.25) | [.66, .81] | .79 (.23) | [.73, .85] | .76 (.24) | [.72, .82] |
| | | | | | | 2/3 Covered | | | | | | |
| Suspect | .57 (.34) | [.47, .68] | .54 (.32) | [.45, .62] | .55 (.33) | [.48, .62] | NA | | NA | | NA | |
| Foil | .16 (.23) | [.09, .23] | .17 (.23) | [.10, .23] | .17 (.23) | [.12, .21] | .39 (.32) | [.29, .49] | .29 (.32) | [.20, .38] | .34 (.32) | [.27, .40] |
| None | .27 (.31) | [.17, .36] | .29 (.29) | [.22, .37] | .28 (.30) | [.22, .34] | .61 (.32) | [.51, .71] | .70 (.32) | [.62, .79] | .66 (.32) | [.60, .73] |
| | | | | | | Fully Covered | | | | | | |
| Suspect | .61 (.26) | [.53, .69] | .53 (.29) | [.45, .61] | .56 (.28) | [.51, .62] | NA | | NA | | NA | |
| Foil | .14 (.20) | [.08, .20] | .16 (.20) | [.10, .21] | .15 (.20) | [.11, .19] | .40 (.32) | [.30, .50] | .31 (.31) | [.23, .40] | .35 (.32) | [.29, .42] |
| None | .25 (.21) | [.18, .32] | .31 (.29) | [.23, .39] | .28 (.26) | [.23, .34] | .60 (.32) | [.50, .70] | .69 (.31) | [.60, .77] | .65 (.32) | [.58, .71] |

*Note*: Means are for the conditions run, rather than the estimated marginal means from conducted analyses of variance. CI = confidence interval; NA = not applicable.