Florida International University FIU Digital Commons

FIU Electronic Theses and Dissertations

University Graduate School

1-11-2013

The Impact of the "Not Sure" Option on Lineup Identification Decisions

Seyram Kekessie Florida International University, skeke001@fiu.edu

DOI: 10.25148/etd.FI13042306 Follow this and additional works at: https://digitalcommons.fiu.edu/etd

Recommended Citation

Kekessie, Seyram, "The Impact of the "Not Sure" Option on Lineup Identification Decisions" (2013). *FIU Electronic Theses and Dissertations*. 850. https://digitalcommons.fiu.edu/etd/850

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

THE IMPACT OF THE "NOT SURE" OPTION ON LINEUP IDENTIFICATION DECISIONS

A thesis submitted in partial fulfillment of

the requirements for the degree of

MASTER OF SCIENCE

in

PSYCHOLOGY

by

Seyram Mawuko Kekessie

To: Dean Kenneth G. Furton College of Arts and Sciences

This thesis, written by Seyram Mawuko Kekessie, and entitled The Impact of the "Not Sure" Option on Lineups Identification Decisions, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Lindsay Malloy

Ronald Fisher

Stephen Charman, Major Professor

Date of Defense: January 11, 2013

The thesis of Seyram Mawuko Kekessie is approved.

Dean Kenneth G. Furton College of Arts and Sciences

Dean Lakshmi N. Reddi University Graduate School

Florida International University, 2013

ABSTRACT OF THE THESIS

THE IMPACT OF THE "NOT SURE" OPTION ON LINEUP IDENTIFICATION DECISIONS

by

Seyram Mawuko Kekessie

Florida International University, 2013

Miami, Florida

Professor Stephen Charman, Major Professor

Law enforcement officials routinely rely on eyewitness identification evidence to solve crimes. Nonetheless, this form of evidence is prone to errors. Researchers have previously attempted to examine conditions under which such errors can be reduced. The present study examines whether giving witnesses an explicit *not sure* response option increases the accuracy of lineup identification decisions. 251 participants watched a mock crime video before viewing a lineup that either included the perpetrator, or was made up of innocent suspects. Results indicated that witnesses provided with a *not sure* option made fewer false identifications, fewer filler identifications, and a similar number of correct identifications as witnesses who were not provided with this option. Furthermore, these benefits occurred regardless of whether witnesses received otherwise biased or unbiased instructions. Results suggest that the inclusion of an explicit *not sure* response option is a simple procedure that can increase the quality of eyewitness lineup decisions.

TABLE OF	CONTENTS
----------	----------

CHAPTER	PAGE
INTRODUCTION	1
Metamemory Reporting	2
Applying Monitoring and Control to a Lineup Identification Task	6
Biased vs Unbiased Lineup Instructions	15
Potential Benefits of the Not Sure Option	16
The Present Study	18
METHOD	19
Participants	19
Materials	19
Procedure	20
RESULTS	21
Not Sure response rate	21
Effects of Not Sure Option on Various Identification Decisions	23
Diagnosticity	
Information Gain Analyses	27
DISCUSSION	
LIST OF REFERENCES	
APPENDIX	40

Introduction

Legal systems in many parts of the world routinely rely on eyewitness identifications as a crucial tool in the pursuit of justice (Weber & Perfect, 2011; Cutler & Penrod, 1995). In western legal systems for instance, witnesses are often presented with a line-up, from which they may attempt an identification (Cutler & Penrod, 1995). Ideally, investigators subsequently use that information to affect an arrest, prosecute a suspect, and serve justice. Except that it does not always end up as seamlessly as expected, and justice may be served erroneously at times. In fact, the identification of innocent suspects has already been identified as a major contributor to injustice in the legal system (Malpass & Devine, 1981). Despite much research documenting the fallibility of eyewitness identification (e.g., Brewer, Keast, & Sauer, 2010; Kohnken & Maass, 1988; Wells & Loftus, 1983; Loftus, 1979; Goldstein, 1977), jurors and judges routinely rely heavily on evewitness evidence when forming opinions and making decisions (e.g., Kohnken & Maass, 1988; Loftus, 1979). The overreliance on eyewitness evidence as a crucial crime-solving tool poses serious challenges to the accuracy of investigation outcomes, and inevitably leads to instances of wrongful conviction of innocent suspects.

Previous research has attempted to address issues raised by the fallibility of eyewitness identification. One approach involves altering the identification process with the goal of enhancing accuracy. Some studies have directly examined conditions under which a witness' identification would be prone to error. For instance, various studies have examined the influence of instruction effects on line-up identification decisions (e.g., Malpass & Devine, 1981; Clark, 2005; Steblay, 1997). The present study also examines the effect of lineup instructions on eyewitness accuracy, but specifically

focuses on the influence of a rarely-researched instruction: the explicit availability of a *not sure* response option. The major goal of the study is to therefore examine whether eyewitnesses' overall performance on a line-up identification task benefits from the inclusion of an option that allows them to decline making a decision if they are unsure, and whether any such benefit is influenced by the type of line-up instructions received.

Metamemory Reporting

Witnesses' decisions rely heavily on their memory processes. These processes may involve the memory or recollection of a particular event, as well as the ability of witnesses to monitor their cognitive processes, a process known as meta-cognition. For instance, when presented with a line-up identification task, witnesses may try to scan the lineup members to determine whether any member matches their memory of a perpetrator. The witness may then make a decision about the likelihood of the match reliably being that of the actual perpetrator. Individuals' ability to monitor their cognitive processes relates closely to the concept of metamemory, which is specifically defined as people's knowledge about their own memory. Understanding metamemory processes would therefore provide valuable insight about how individuals' cognitive processes influence the recall and reporting of information. In recent years, metamemory has become an important research area enabling better understanding about how individuals make decisions, and the processes that lead to those decisions.

In a series of experiments, Koriat and Goldsmith (1996) proposed a metamemory reporting model, which has proven quite useful as a solid overarching theoretical framework for examining metamemory. Koriat and Goldsmith's (1996) model builds on Nelson's (1996) definition of metacognition as consisting of two processes. The first

process is monitoring, which refers to an individual's gathering of information about his knowledge (Spehn & Reder, 2000). Memory performance is influenced by the degree to which individuals can access their memory and match information to their latent memory trace.

Koriat and Goldsmith (1996) however suggest that memory performance goes beyond simply retrieving information from memory. In addition, memory performance is a function of the degree to which individuals monitor the probability of correctness of any range of responses that come to mind, and choose a response accordingly. For example, the individual can decide on the exact details to divulge, the ones to withhold, the ones he would emphasize, and so forth (Koriat & Goldsmith, 1996; Ross & Buehler, 1994), a process known as control. Control refers to the process of self-regulation of memory responses (Spehn & Reder, 2000). During the control process, the individual monitors the likelihood of correctness of the range of responses that come to mind, and selects a response based on a self-determined decision criterion or control threshold. Essentially, individuals weigh the response with the highest subjective probability of correctness against the threshold, and if the probability passes the threshold, that particular response is reported. Conversely, if it does not pass the threshold, the individual may choose to withhold the response. The criterion for making a response (decision criterion) can be adjusted on the basis of costs or benefits for reporting accurate or inaccurate responses (Rhodes & Kelley, 2005). For instance, a strong motivation to be accurate would raise the decision criterion and result in high selectivity in reporting responses, thereby resulting in generally greater levels of memory accuracy. Conversely,

a lower motivation to be accurate would result in low selectivity and therefore generally lesser levels of accuracy.

In more concrete terms, consider an example where the question is "what is the capital of Germany?" Using a similar example, Koriat and Goldsmith (1996) suggest that participants would have in mind a range of responses such as "Berlin," "Hamburg," and "Munich." They then assign a probability of correctness (Pc) for each of the responses, and subsequently set a self-determined decision criterion (Dc) in order to control their responding. Assuming the Pc for "Berlin," "Hamburg," and "Munich" is .70, .40, and .35 respectively, and the Dc is set at .60 probability of correctness, participants would choose to report the response with a Pc equal or greater than .60, and withhold the response that falls short of the decision criterion. Consequently, "Berlin" would be offered as the response in this example, and "Hamburg" and "Munich" would be withheld.

Koriat and Goldsmith's (1996) model permits an examination of memory accuracy-based measures of performance within the traditional memory assessment framework. The complex, systemic interplay of monitoring and control processes influences the probability that a reported response is accurate. Accordingly, Koriat and Goldsmith predict that allowing witnesses to regulate their memory reporting would lead to greater accuracy levels in memory performance. Witnesses regulate their reporting by withholding responses for which they are unsure, and reporting the ones they perceive to have a high probability of correctness. Specifically, the model would predict that providing witnesses with an explicit *don't know* or *not sure* response option will allow witnesses to opt out of making responses when they are uncertain, which should in turn lead to high memory performance accuracy. In essence, people who would have

otherwise given a response with low certainty would be able to opt for an explicit *don't know* or *not sure* response option, thereby reducing the likelihood of giving an incorrect response. For the purposes of the present study, it is assumed that a *don't know* option is essentially equivalent to a *not sure* option, since they both refer to uncertainty about making a response. Therefore, these terms are used interchangeably, as has been the case with previous research on this topic (e.g., Koriat and Goldsmith (1996) manipulate *don't know*, whereas Brewer, Keast and Sauer (2010) manipulate *not sure*).

To evaluate the prediction that the option to withhold responses would lead to higher accuracy in memory performance, Koriat and Goldsmith (1996) presented a group of participants with a set of general knowledge questions and forced them to respond (forced-report condition). These participants were subsequently asked to give a confidence judgment for each answer. The participants' answers and their confidence judgments allowed experimenters to estimate a participant's correct response probability for each question. A second group of participants were asked similar general knowledge questions, but they were not required to give a response (free-report condition). In effect, a participant could withhold a response when unsure of the answer to the question.

The results indicate that allowing individuals to control their response output, by providing a *don't know* option which allows them to opt out of making unsure responses, leads to an improvement in accuracy. Forcing witnesses to guess compels them to provide responses that did not meet their decision criterion, which tend to be more inaccurate. Allowing them a *don't know* option allowed those witnesses to opt out of responding if the Pc of a response option didn't surpass their decision criterion. Taken as

a whole, the results revealed greater accuracy for reported items in free-report conditions than forced-report conditions.

Koriat and Goldsmith's (1996) findings suggest that the memory accuracy accompanying the utilization of the *don't know* option came at no significant cost to the number of correct responses. According to Koriat and Goldsmith, because people are able to effectively monitor their memory, they would be able to distinguish between correct and incorrect answers and respond accordingly. In essence, a perfect monitoring system, coupled with an effective control system would equate to excellent memory performance.

It is important to note, however, that Koriat and Goldsmith tested participants' knowledge of general knowledge information, rather than their memory of faces. Their finding may therefore be explained by the possibility that people have particularly strong metacognitive abilities when it comes to general obscure trivia. On the other hand, people may have weaker metacognitive abilities when dealing with imperfect memories of faces, hence the need to evaluate this model in the context of a facial identification task.

Applying Monitoring and Control to a Line-up Identification Task

Eyewitness identification decisions in particular are influenced by multiple factors, including the accuracy of the witness's memory trace and the willingness to release or withhold a decision depending on the individual's assessment of his own memory (Clarke, Howell, & Davey, 2008). Koriat and Goldsmith's work provides a framework within which witnesses' decision-making processes on a line-up identification task can be evaluated. The complex interplay of monitoring and control processes influences the decision-making processes of witnesses. For instance, when presented with a simultaneous line-up, witnesses are able to scan the line-up members simultaneously.

The monitoring process begins at this stage when witnesses compare each lineup member to their memory trace. Assuming the witness has an adequate memory trace, the perpetrator in a target-present (TP) line-up should tend to produce the greatest Pc value (i.e., he should be the best match to the witnesses' memory trace). Conversely, lineup members in target-absent (TA) line-ups should tend to be relatively poor matches to memory (i.e., they should produce relatively low Pc values). Of course, the actual response witnesses report or choose to withhold depends on the decision criterion they set during the control process.

The decision criterion, which refers to witnesses' willingness to make an identification decision, is influenced by factors such as the relative costs and benefits for inaccurate and accurate decisions. Witnesses' decision criteria may have different implications for their memory performance in TA and TP line-ups. For instance, an accuracy-motivated witness with a high decision criterion (i.e., a high control threshold for response) would be highly selective with responses and could therefore perform strongly (i.e., make a correct identification) when shown a TP line-up since there is one member of the line-up (the suspect) who strongly matches his memory trace (again, assuming a strong memory trace). However, a high decision criterion does not always translate into strong memory performance since memory traces are not infallible. Consequently, a high decision criterion can lead a witness to mistakenly fail to identify the perpetrator from a TP line-up. Nonetheless, a higher decision criterion will tend to screen out identifications of lineup members that do not produce strong recognition experiences, which should tend to be innocent people or fillers. Similarly, witnesses viewing a TA line-up could correctly monitor their memory, recognize that none of the

line-up members match their memory trace of the suspect, and consequently withhold potentially incorrect responses or correctly reject the line-up.

It is worth examining how the Koriat & Goldsmith model applies specifically to a *not sure* option in a lineup task. In theoretical terms, the *not sure* option may have the effect of increasing the decision criterion (Koriat & Goldsmith, 1996), as it allows witnesses to screen out unsure responses, thereby increasing selectivity in responding. The overall effect therefore should be a reduction in false identifications, and a corresponding improvement in the quality of memory output. There is a possibility of a drop-off in hits as well, but because the Pc of the perpetrator will be higher on average than the Pc of an innocent person, a not sure response should decrease false alarms more than hits.

To be sure, traditional line-up research procedure has usually provided witnesses with a *not there* option to allow for a rejection of the line-up. Although a *not there* option may at first glance appear to provide the same function as a *not sure* option, Koriat and Goldsmith's (1996) metamemory framework suggests that they are conceptually different. A *not there* response indicates that witness has actually evaluated the line-up members and has concluded that the actual perpetrator is "not there." In contrast, a *not sure* option allows witnesses to screen out responses for which they are uncertain and which may be potentially inaccurate.

For instance, according to Clark's (2003) WITNESS model, a mathematical model describing witness choosing behavior from lineups, witnesses have two decision criteria – a criterion for acceptance and a criterion for rejection. If the evidence against a lineup member (a combination of the match to memory produced by that lineup member,

and the difference between the match produced by that lineup member and the match produced by the next best lineup member) is above the criterion for acceptance, the witness will make an identification. If, on the other hand, it is below the criterion for rejection, a 'not there' response will be given. But what occurs when the match value or Pc is between the two criteria? In that case, the model predicts that witnesses will opt for a *not sure* response. Consequently, a *not there* response option is conceptually different from a *not sure* response option. Thus, the scarcity of research that directly manipulates the availability of a *not sure* option to examine potential benefits to witnesses' memory accuracy provides justification for the present study.

Findings from eyewitness studies tend to show a drop off in hits when the decision criterion is raised. For instance, biased line-up instructions lead witnesses to raise their decision criterion, which in turn increases selectivity in responses, reducing both false as well as correct identifications (Clark, 2005). Similarly, sequential lineups, in which lineup members are shown one at a time to the witness – tend to result in fewer false identifications and fewer correct identifications, consistent with the idea that sequential lineups raise a witness's decision criterion (Steblay, 2001; Ebbeson and Flowe, 2002).

Despite the drop-off in hits, raising a decision criterion tends to lead to at least an equal, if not greater, drop-off in false identifications compared to hits. For instance, while some studies involving the *not sure* option found an equal drop-off in false and correct identifications, Other studies have even found a greater drop off for false identifications than correct identifications, which can be ascertained via diagnosticity measures; Wells & Olson (2002).

Diagnosticity of a response simply refers to the degree to which the response reveals information about guilt or innocence, which in turn allows for an updating of one's beliefs in the guilt/innocence of a suspect (Wells & Olson, 2002; Wells & Lindsay, 1980). Diagnosticity estimates of a suspect identification are derived by calculating the ratio of guilty suspect identifications to innocent suspect identifications (across witnesses). The issue of diagnosticity was investigated by a meta-analytic study examining the diagnosticity of identifications in both target-absent (TA) and targetpresent (TP) line-ups. Clark, Howell, and Davey (2008) analyzed comparisons of studies that included the "don't know" option for identifications. Their analyses found that don't know responses were entirely non-diagnostic and provided no information about guilt or innocence, strongly implying that these responses merely undermined the diagnosticity of the other responses. Clark, Howell, and Davey's (2008) finding contradicts other studies (e.g., Wells & Olson, 2002), which found that don't know responses were slightly diagnostic of a suspect's innocence. However, it is difficult to generalize from Clake, Howell, and Davey's (2008) results since the studies they used in their comparisons did not experimentally manipulate the availability of the *don't know* option

A more important question is whether a *not sure* option affects accuracy. Indeed, only a few line-up studies (e.g., Brewer, Keast, & Sauer, 2010; Steblay & Philips, 2011; Weber & Perfect, 2011) have directly manipulated the availability of the *don't know* or *not sure* option to examine its impact on witness identification accuracy using Koriat and Goldsmith's (1996) model as a theoretical framework. In Brewer, Keast and Sauer (2010), participants (children) viewed video clips of simulated crimes varying in terms of crime nature and exposure time. After viewing the stimulus video, participants were

randomly assigned to one of three instruction conditions and administered a lineup. All participants in the study were administered unbiased lineup instructions. Participants in the *not sure* condition were told to click on the *not sure* button if they "don't know or are not sure that the thief is in the lineup". They were also told that it was "better to click on the *not sure* button than to guess." In the *accuracy motivation* condition, participants were given manipulations designed to motivate them to be accurate, in addition to a *not sure* response option. They were encouraged to be accurate through the awarding of points: 10 points for a correct identification or correct rejection, 5 points if they clicked *not sure*, and 0 points for an incorrect rejection or false identification. In the control condition, the participants were not given any accuracy motivation manipulations or a *not sure* response option.

The results indicated that the availability of the *not sure* option did not impact the false identification or target identification rate. However, the accuracy motivation conditions coupled with the not sure option succeeded in reducing guessing, with the overall rate of false identifications dropping by approximately 12%. It is important to note the *not sure* option did not significantly reduce hits, thereby suggesting that the *not sure* response option at the very least does not impair performance. In contrast, it has the potential to improve performance, at least if coupled with a motivation instruction. These results about the benefits of the *not sure* option can be described as modest at best. However, it is worth noting that the study sampled children around the age of 12. Accordingly, it is difficult to determine the extent to which these results can be generalized, given that adults or even older children may respond differently to lineups.

For instance, children have been shown to guess significantly more than adults on identification tasks (Pozzulo & Lindsay, 1998).

Using an adult sample, another study formally tested the effects of the *don't know* option by experimentally manipulating its availability to witnesses presented with an identification task. Weber and Perfect (2011) describe an approach that contrasts the availability of an explicit *don't know* option with standard unbiased line-up instructions. After viewing a crime video and a subsequent filler video, participants were randomly assigned to one of three response option conditions and a show-up was administered. In the forced-report condition, participants were instructed to respond *yes* if the photograph was that of the offender, and *no* if otherwise. In the explicit free-report condition, participants were provided an additional option allowing them to give a *don't know* response if they were unsure. Finally, those in the spontaneous free-report condition were not provided with any response options. Instead, they were told to type their response in their own words, and were free to use whatever words to indicate their response. The results indicated that the not sure option reduced false identifications without a concomitant decrease in correct identifications, thereby increasing the diagnosticity of identifications

While Weber and Perfect's (2011) results are consistent with prior studies showing an accuracy advantage for witnesses allowed the option of withholding or reporting a response, the study has some limitations. Firstly, the choice of using a showup instead of a line-up can impact the overall generalizability of their findings, given that line-ups are considered inherently less suggestive than show-ups (Wells & Quinlivan, 2009). Compared to show-ups, line-up decisions are more complex and can involve

exposure to different stimuli simultaneously. Specifically, show-ups involve direct comparison of the suspect and memory, whereas line-ups involve comparisons among lineup members (Wells, 1993). In line-ups, witnesses have an opportunity to simultaneously view the line-up members, including foils, and make a decision on the basis of the perceived similarity to their memory match of the perpetrator, and their decision criterion. The process is best illustrated with Clark's (2003) WITNESS model, which suggests that witnesses presented with a simultaneous lineup engage in relative judgments to determine the best-next match difference. Show-ups in contrast involve a single suspect, who the witness must either reject or identify without being able to compare with other members who may or may not be similar matches to their memory trace. Consequently, it is reasonable to expect that cognitive processes from a lineup will differ from a show-up.

Furthermore, there is applied importance to understanding the extent to which *don't know* responses come from people who would have otherwise made filler identifications as opposed to suspect identifications. Show-ups in contrast preclude the possibility of examining whether *don't know* responses are provided by people who would have otherwise made filler identifications, since a show-up by its nature does not have fillers. In show-ups, any decrease in the choosing rate would necessarily translate to a reduction in target identifications. In lineups, however, a decrease in the choosing rate could either mean a decrease in target identifications, or rather filler identifications. Because a filler is a known innocent person, unlike an innocent suspect, these two types of identifications are conceptually distinct, and it is thus important to examine the effect of a not sure option on lineup decisions.

In a recent study, Steblay and Philips (2011) examined the influence of a *not sure* option on witness accuracy in simultaneous and sequential lineups. Participants viewed a 30-second crime video of a purse snatching incident, and subsequently attempted to identify the perpetrator from either a simultaneous or a sequential lineup. Witnesses in the forced-choice condition were provided with a dichotomous (*yes/no*) response choice, and those in the *not sure* condition were provided an explicit *not sure* response option in addition to the dichotomous (*yes/no*) response choices. All participants were given unbiased lineup instructions, and randomly assigned to view either a simultaneous or sequential lineup. The *not sure* option led to a decrease in false identifications as well as correct identifications for witnesses who viewed a sequential lineup. In contrast, the *not sure* option had minimal impact in the simultaneous lineup; it failed to significantly reduce correct identifications or false identifications.

The results indicate difficulties with generalizing the benefits of a *not sure* option across different lineup formats, since the *not sure* option had minimal impact on simultaneous lineups. One reason for this may be due to the strength of the manipulation of the *not sure* option, given that the *not sure* option was manipulated only in the form of a response option. In other words, participants were not instructed prior to viewing the lineup that they had an option of saying *not sure*. The present study used a stronger manipulation procedure by presenting the *not sure* option both as a response option and as an instruction (witnesses in the not sure condition were instructed prior to viewing the lineup that they could respond *not sure* if they were uncertain). Additionally, Steblay and Philip's (2011) study is one of very few studies directly examining the influence of the

not sure option using lineups specifically, and therefore there is a need to replicate and expand upon their findings.

Steblay and Philips's (2011) results with regards to sequential lineups are in line with previous line-up research showing that a high decision criterion tends to reduce hits (e.g., Clark, 2005; Ebbeson & Flowe, 2002), assuming that the *not sure* option has the effect of increasing the decision criterion. However, results from other studies (e.g., Weber & Perfect, 2011) lead to contrary conclusions. Given the conflict in these findings, further research is needed to further examine the effectiveness of the *not sure* option. The present study will examine the influence of the *not sure* option on witnesses' identifications. Similar to prior studies, it is expected that the *not sure* option would have the effect of improving memory accuracy by reducing false identifications more than it reduces correct identifications. However, this effect may be moderated by a different type of line-up instruction.

Biased vs Unbiased Line-up Instructions

Previous research has established that identification accuracy depends on line-up instructions – specifically, whether the lineup administrator explicitly tells witnesses that they are allowed to indicate that the perpetrator is not present in the lineup (Steblay, 1997). Biased line-up instructions that do not allow for a 'not there' response are highly suggestive of the presence of the perpetrator in the line-up and discourage a "no choice" response. Previous research (e.g., Clark, 2005; Clarke, Howell, & Davey, 2008) has established that biased instructions generally lead to a greater willingness to make an identification than do unbiased instructions. In the context of the monitoring and control framework, biased instructions produce a lower decision criterion than unbiased

instructions. The explicit suggestion of the presence of the perpetrator in the lineup results in witnesses who are more willing to choose, and therefore less selective in their responses, compared to unbiased instructions. The overall effect consequently is that biased instructions lead to more false identifications (and slightly more correct identifications) than unbiased instructions (Clark, 2005; Steblay, 1997).

The aforementioned studies examining the influence of the not sure or don't know option on identification accuracy solely utilized unbiased line-up instructions. However, there may be strong reasons why the impact of the not sure option could differ across line-up instruction type. Specifically, the not sure option may be more beneficial when given alongside biased instructions rather than unbiased instructions. In the presence of biased instructions, the not sure option may raise witnesses' decision criteria and consequently increase their selectivity in responding, which would in turn lead to greater levels of accuracy. Under unbiased instructions, however, the decision criterion is already high and witnesses may already be highly selective in responding. Therefore, an already high decision criterion will result in ceiling effects, blunting the effectiveness of the not sure option. Consequently, it is unclear whether the not sure option would improve accuracy under unbiased line-up instructions to the same degree that it would under biased instructions. Thus, if the beneficial effect of the not sure option is redundant with unbiased lineup instructions, it may not add any additional diagnostic value. On the other hand, if it is not redundant, it can be recommended for use in addition to unbiased instructions.

Potential Benefits of the Not Sure Option

There is a strong rationale for studying the effects of allowing witnesses the option to withhold or report memory responses through the provision of an explicit *not sure* option. First, the provision of an explicit *not sure* option enables witnesses who are unsure about their responses to indicate their doubts about the reliability of their decision. One may argue that alternatively, low confidence witnesses may simply make a decision, and rate their confidence in the identification as low, thus allowing triers of fact to screen out inaccurate identifications via witnesses' low confidence, and that consequently a *not sure* option is unecessary. However, research has shown that witnesses' initial low confidence can get inflated over time via factors such as repeated questioning and feedback (e.g., Shaw, 1996; Wells & Bradfield, 1998). Filtering out low confidence witnesses into *not sure* responses would prevent that confidence inflation problem from ever occurring in the first place. Allowing witnesses to withhold an unsure response may be a more effective procedure than relying solely on confidence as an indicator of accuracy.

The second benefit of the *not sure* option is that it may potentially reduce false rejections from target present lineups. A *not there* response, if given with low confidence, may be inaccurate. However, a *not sure* response when available may be used by witnesses who otherwise would have falsely rejected the lineup. Clark's (2003) WITNESS model is helpful in illustrating this point. If the *not sure* option only increases the decision criterion (as suggested by Koriat and Goldsmith's model), then it should not lead to a decrease in false rejections. However, the WITNESS model, which states that there are two criteria, predicts that the availability of a *not sure* option can have separate

effects on the two criteria, increasing the criterion for making an identification and simultaneously reducing the criterion for rejection independently. In other words, both low confident identifiers and low confident non-identifiers will both tend to shift to not sure responses. This consequently enables some non-choosers who would have otherwise falsely rejected the lineup to opt for a *not sure* response, potentially leading to a decrease in mistaken lineup rejections. If true, this would be a welcome benefit, as very few studies have examined ways to reduce false rejections.

Along a similar line, the third benefit of the *not sure* option is the potential to reduce filler identifications if those prone to making filler identifications are low confidence witnesses who would be screened out if the *not sure* response option is available. This is very important since identifying a filler will undermine a witness' credibility if subsequently presented with a second lineup.

As already discussed, some previous research (e.g., Weber & Perfect, 2011) has found that the *not sure* option reduces false identifications. The relevant practical question, however, is the extent to which the reduction of false identifications comes at a significant cost to the number of correct identifications. The usefulness of allowing witnesses the option to withhold unsure responses will depend on the extent to which it reduces incorrect identifications without significantly undermining the amount of correct identifications.

The Present Study

The present study experimentally manipulates line-up instruction type to examine the benefits of a *not sure* option under conditions of biased and unbiased instructions. By providing an explicit *not sure* option, this study examines whether witnesses are able to

withhold making a decision when they have doubts about the likely accuracy of that decision. This should influence overall accuracy. The following predictions are made: the availability of a *not sure* option will (a) lead to a reduction in filler identifications, (b) lead to a reduction in false rejections from target present lineups, (c) lead to a reduction in false identifications from target absent line-ups, and (d) lead to a smaller reduction in correct identifications from target present lineups; but (e) these effects will be moderated by instruction type: They should be greater when witnesses are given biased, as opposed to unbiased, instructions.

Method

Participants

Two hundred and fifty one undergraduate students at Florida International University participated in the study. All participants were recruited through SONA systems, and received extra credit for their participation in the study.

Materials

Mock crime video. The stimulus video clip shows a staged crime in which the offender steals an item from a victim's purse. The offender was seen approaching the purse situated on a bench, and he proceeded to pick an item from the purse, and fled the scene. As he fled, the perpetrator's face was in direct frontal view for about 5 seconds. The video was presented on 17-inch desktop computer monitors with a resolution of 1280 x 1024.

Lineups. To select the photographs used in constructing the lineups, 4 research assistants were asked to watch the video and provide specific descriptions of the characteristics of the perpetrator. The five most common description items were "male,"

"dark hair," "Hispanic," "tall," and "average build." This list was subsequently given to a different set of 4 research assistants, who looked through a set of 20 photographs for description matches. The seven most selected photographs (including the photograph of the actual perpetrator) were used to construct the lineups. Two types of line-ups were created. Target absent lineups consisted of six members (fillers), none of which was the perpetrator. These fillers matched the description of the perpetrator. The target present line-up consisted of a six-member line-up, with the perpetrator included as a member of the line-up. Six different target present lineups were created in total, with the perpetrator replacing a different member of the target-absent lineup in turn. All the images were cropped into a mug shot of the line-up member, revealing only the head and shoulders against a white background.

Filler video. Participants were shown a brief video clip during the interval between when they viewed the stimulus video and when they made identification decisions. The video was specifically a 3-minute long nature video that involved different animals in their various natural habitats.

Procedure

Upon arrival, each participant was assigned a computer on which all stimuli were presented via Microsoft PowerPoint. The basic procedure, which all participants followed, entailed watching the stimulus video first, followed by the non-related filler video clip, and then being presented with a randomly assigned line-up and asked to make an identification decision. There were three factors manipulated between-subjects: the nature of the line-up (target absent or target present), the presence or absence of the 'not sure' response option, and line-up instructions (biased or unbiased).

After watching the stimulus video clip, participants viewed the filler video clip. Immediately after watching the filler video clip, they were given condition-specific instructions. Participants were randomly assigned to: (a) either biased or unbiased lineup instructions (i.e., will be given the option of responding *not there* or will not be given the option of responding *not there*), (b) either a *not sure* option or not (i.e., will be given the option of responding *not sure* or will not be given the option of responding *not sure*), and (c) either a target absent or target present line-up. See Appendix A for the complete set of instructions given to participants in each condition. After providing their identification responses, all witnesses were given a questionnaire assessing their confidence and other testimony-relevant measures. Lastly, participants provided demographic information.

Results

Overview

First, the frequency of *not sure* responses in general was examined in target absent and target present lineups, when lineup instructions are either biased or unbiased. Secondly, witnesses' identification decisions were examined across the various conditions to determine how the presence of a *not sure* option affected identification decisions (suspect identifications, filler identifications, and *not there* responses). Third, analyses were conducted to determine the diagnosticity of witnesses' responses, including *not sure* responses. Fourth, information gain analyses were conducted to examine the informative value of responses. An alpha level of .05 was adopted for all analyses, and *Phi* (ϕ) was reported as a measure of the effect size. The cutoff for small effects was set at .1, medium effects at .3, and large effects at .5.

Not Sure Response Rate

One of the first questions examined is whether participants would withhold a response when given the option to do so. Accordingly, analyses were conducted to specifically compare the frequency of *not sure* responses in target present and target absent lineups, when lineup instructions were biased versus unbiased (see Table 1 for descriptive statistics). For target present lineups, a 2 (Lineup Instruction: biased vs. unbiased) x 2 (Witness Decision: withhold (*not sure*) vs. report (non-*not sure* response) Chi square test was performed to determine the extent to which the frequency of *not sure* responses differed across lineup instruction. The results revealed no significant difference between the frequency of *not sure* responses produced when lineup instructions were biased and when lineup instructions were unbiased, χ^2 (1) = .19, p = .67, φ = .04. For target absent lineups, a 2 (Lineup Instruction: biased vs. unbiased) x 2 (Witness Decision: withhold (*not sure*) vs. report (non-*not sure* response) and the lineup instructions were unbiased, χ^2 (1) = .19, p = .67, φ = .04. For target absent lineups, a 2 (Lineup Instruction: biased vs. unbiased) x 2 (Witness Decision: withhold (*not sure*) vs. report (non-*not sure* response) Chi square test revealed that *not sure* responses were provided more often when lineup instructions were biased, as compared to when lineup instructions were unbiased, χ^2 (1) = 5.19, p = .02, φ = .2.

Furthermore, the results revealed that participants generally produced *not sure* responses at a greater frequency when presented with a target absent lineup than when presented with a target present lineup, $\chi^2(1) = 17.27$, p < .01, $\varphi = .37$. As illustrated in Table 1, 11.1% of people presented with TP lineups made a not sure response, compared to 43.9% of people presented with TA lineups.

Overall, 28% of participants opted to make a *not sure* response when the option was available. However, it is difficult to infer accuracy from solely examining the frequency of *not sure* responses. For instance, it is difficult to conclude whether *not sure*

responses came from people who would have otherwise produced correct identifications or rejections. Subsequently, there is a need to examine the extent to which the ability to withhold a response influences the suspect identification rate, the filler identification rate, and the lineup rejection rate separately.

Effects of not sure option on various identification decisions

Target present lineups. Collapsing across lineup instructions, a 2 (Not sure Option: yes vs. no) x 2 (Target ID: accurate vs. inaccurate) Chi square test was performed to examine whether memory performance accuracy was influenced by the availability of the option to withhold a response. The availability of a *not sure* option did not significantly influence the likelihood of making a correct identification, $\chi^2(1) = .11$, p = .74, $\varphi = .03$. (See Table 2 for descriptive statistics).

Interestingly, the availability of the *not sure* option did influence the filler identification rate from target present lineups. Collapsing across biased and unbiased lineup instructions, a 2 (Not sure Option) x 2 (Filler ID) Chi square test revealed that the availability of the *not sure* option decreased filler identifications, χ^2 (1) = 4.04, p = .05, φ = .18. A witness was approximately twice as likely to make a filler ID when the *not sure* option was not available (31%) than when it was available (16%).

Collapsing across lineup instruction, a 2 (Not sure Option) x 2 (False Rejection) Chi square test was performed to examine the influence of a *not sure* option on false rejections in target present lineups. The availability of a *not sure* option did not significantly influence the false rejection rate, $\chi^2(1) = .04$, p = .85, $\varphi = .02$.

Target absent lineups. In TA lineups, any identification of a lineup member is by definition inaccurate. In the present study, the suspect versus filler role was rotated

among all the six line up members, in lieu of designating an a-priori suspect. Thus, the false ID rate is calculated by dividing by 6 the number of participants who identified any member of the lineup (that is, any response except a *not there* or *not sure* response). Previous eyewitness studies have adopted this procedure for analysis (e.g., Charman & Wells, 2007). *Not sure* responses were not considered inaccurate for the purpose of this analysis.

Among participants who received biased lineup instructions, a 2 (Not sure Option) x 2 (Choosing: chooser vs. non-chooser) Chi square test revealed that the availability of the *not sure* option decreased the choosing rate, χ^2 (1) = 25.67, p = .01, φ = .63. As shown in Table 3, 100% of participants falsely identified a lineup member as the perpetrator when there was no option allowing them to withhold a response. However, the choosing percentage dropped to 41.2% when the *not sure* option made available. Thus, the false identification rate was reduced from 16.67% when the *not sure* option was not available, to 6.87% when it was available. This result is hardly surprising, since participants who are not allowed the option of withholding a response are basically forced to identify (falsely) a member of the lineup.

A 2 (Not sure Option) x 2 (Choosing) Chi square test was performed among participants who received unbiased instructions. The results revealed significant main effects of the availability of the *not sure* option on the choosing rate, such that the *not sure* option decreased the choosing rate, $\chi^2(1) = 3.84$, p = .05, $\varphi = .25$. As shown in Table 3, the choosing rate is reduced from 45.2% when the *not sure* option was not available, to 21.9% when it is available. Thus, the false ID rate falls from 7.5% when the *not sure* option is not available to 3.7% when the *not sure* option is available.

It is potentially misleading to examine whether the effect of the *not sure* option on choosing is moderated by biased vs unbiased instructions, because the choosing rate in the biased instruction condition (100%) was very discrepant from the choosing rate in the unbiased condition (45.2%), thus potentially producing differential floor effects. It may be more meaningful to examine instead the *proportional* reduction in false identifications across conditions. This proportional reduction was similar for witnesses who received biased lineup instructions and those who received unbiased lineup instructions: Availability of the *not sure* option led to a 58.8% decrease in false identifications for witnesses who received biased lineup instructions, compared to a 51.5% decrease in false identifications for witnesses who received unbiased lineup instructions, suggesting that the effect of a *not sure* option was not dependent on whether the lineup instructions were biased.

The filler identification rate in target absent lineups is calculated as five-sixths of the choosing rate (since the other one-sixth are considered suspect identifications). Collapsing across lineup instructions, there was a reduction in filler identifications in target absent lineups from 60% when the *not sure* option was not available, to 27% when it was available. Taken together, these results mean that the presence of a *not sure* option not only serves to reduce false identifications, but also reduces filler identifications as well, while avoiding a drop off in the correct identification rate.

Diagnosticity

In eyewitness lineup research, diagnosticity of a response is calculated to determine the extent to which that particular response provides information about a suspect's guilt or innocence. Specifically, diagnosticity estimates provide information

about the likelihood that an identified suspect is the actual perpetrator, rather than an innocent person (Wells & Lindsay, 1980; Steblay & Philips, 2011). This is helpful when comparing different lineup procedures to determine which is superior in probative value (Wells & Olson, 2002; Charman & Wells, 2007). Diagnosticity of an incriminating response (i.e., a suspect identifications) is calculated as the ratio of the likelihood of a particular response in target present lineups to the likelihood of that response in target absent lineups. Diagnosticity of an exonerating response (i.e., a filler identification, a not there response, or a *not sure* response) is calculated as the ratio of the likelihood of a particular response in target absent lineups to the likelihood of that response in target present lineups. The greater the diagnosticity, the more information it provides about the guilt or innocence of a suspect. First, the diagnosticity of *not sure* responses was calculated for participants who received biased and unbiased lineup instructions. Subsequently, the diagnosticity of target identifications was calculated for participants who received biased and unbiased lineup instructions, when the *not sure* response option was available as well as when it was not available. Inferential statistics for these measures (i.e., whether diagnosticity differed significantly from zero) are provided in the 'Information gain analyses' section.

Diagnosticity of *not sure* **responses.** Among those who received biased lineup instructions, diagnosticity of a *not sure* response was 4.42, meaning that a witness was 4.42 times more likely to respond *not sure* when the suspect is innocent than when guilty. For those who received unbiased lineup instructions, the diagnosticity of a *not sure* response was 3.09, meaning that a witness was 3.09 times more likely to respond *not sure* when the suspect is innocent than when guilty.

Diagnosticity of target identifications. To calculate diagnosticity estimates, the percentage of identifications (choosing rate) in TA lineups was divided by six to obtain the average false identification rate.

Among participants who received biased lineup instructions, the availability of a *not sure* option increased the diagnosticity of a correct identification from 3.29 to 8.73. Among participants who received unbiased lineup instructions, the availability of a not sure option increased the diagnosticity of a correct identification from 7.52 to 15.78.

Information gain analyses

Information gain analyses enable us to examine the extent to which we can update our knowledge about the guilt of a suspect, based on a witness' response. The information gained from a lineup decision depends on (a) the diagnosticity of that decision, and (b) the likelihood of a suspect's guilt, prior to knowing the witness' performance on an identification task (Wells & Olson, 2002). It is derived by examining the difference between the prior probability (the likelihood that the suspect is the actual culprit before the line was administered) and the posterior probability (the likelihood that the suspect is the actual culprit after the lineup is administered and the witness has provided a response). Peak information gain for an identification occurs at relatively low levels of the prior probability (i.e., less than p = .5), because new information has a greater impact if it is counter to preexisting beliefs than if it is consistent with preexisting beliefs (Wells & Olson, 2002). Similarly, peak information gain for a not sure response should peak at relatively high levels of the prior probability (i.e., greater than p = .5) for the same reason. In the present study, information gain was calculated using Bayesian analyses, and the information gain was graphed (see Figure 1, Figure 2, and Figure 3) as

a function of the suspect's prior probability of guilt (Wells & Olson, 2002; Charman & Cahill, 2012). The amount of information gained was calculated for both target identifications as well as *not sure* responses associated with the present data.

Information gain for *not sure* **responses.** Among witnesses who received biased lineup instructions, information gain associated with the knowledge that a witness made a *not sure* response peaked at approximately 0.35 (at a prior probability of approximately 0.7), and was significantly greater than zero, z = 4.04, p = < 0.001. Among witnesses who received unbiased lineup instructions, information gain associated with a *not sure* response peaked at approximately 0.27 (at a prior probability of approximately 0.65), and was significantly greater than zero, z = 2.00, p = 0.02. These results indicate that there is significant information gained by knowing that the witness made a *not sure* response; specifically, it tends to provide information that the suspect is innocent. A test examining whether the curve associated with a *not sure* response when lineup instructions are biased differed from the curve associated with a *not sure* response when lineup instructions were unbiased revealed a marginally significant difference between the two curves, z = 1.42, p = 0.08.

Information gain for target identifications. Among witnesses who received biased lineup instructions, Information gain was calculated separately for target identifications when the *not sure* response option was available and for target identifications when the *not sure* response was not available. When the *not sure* option was not available, information gain associated with the knowledge that a witness correctly identified the suspect peaked at approximately 0.29 (at a prior probability of approximately 0.35), and was significantly greater than zero, z = 3.15, p = < 0.001. In

comparison, when the *not sure* option was available, information gain associated with a suspect identification peaked at approximately 0.49 (at a prior probability of approximately 0.25), and was significantly greater than zero, z = 4.92, p = < 0.001.

The results reveal that there is significant information gained by knowing that a witness correctly identified the suspect. Subsequent analyses revealed that, among witnesses who received biased lineup instructions, the availability of a *not sure* option did not significantly change the amount of information gained from a suspect identification, z = 1.14, p = 0.13.

Among witnesses who received unbiased lineup instructions, suspect identifications when the *not sure* response option was not available peaked at approximately 0.46 (at a prior probability of approximately 0.25), and was significantly greater than zero, z = 4.38, p = < 0.001. In comparison, when the *not sure* option was available, information gain associated with a suspect identification peaked at approximately 0.59 (at a prior probability of approximately 0.2), and was significantly greater than zero, z = 5.36, p = < 0.001.

Subsequent analyses revealed that, among witnesses who received unbiased lineup instructions, the availability of a *not sure* option did not significantly change the amount of information gained from a suspect identification, z = 0.56, p = 0.29.

Discussion

The present study explored whether the availability of a *not sure* option influenced witnesses' performance on a lineup identification task, and whether that would be moderated by biased versus unbiased lineup instructions. A key underlying question is whether providing an explicit *not sure* option would increase overall accuracy. The results supported the hypothesis that allowing witnesses the option to withhold a response would lead to a reduction in false identifications. This suggests that the *not sure* option provided an avenue for witnesses to withhold a response when in doubt about the likelihood of correctness of an identification. Many witnesses who made *not sure* responses would have otherwise made a false identification in the absence of the option to withhold responding. This effect of the *not sure* option is significant and important for the reason that reducing false identifications reduces the likelihood of prosecuting an innocent suspect.

However, there was the possibility that providing a not sure option might reduce hits as well as false alarms, resulting in no net improvement in overall accuracy, or even a net cost to accuracy. The underlying reasoning is that if the not sure option is supposed to improve memory accuracy by screening out answers that are perceived to be likely inaccurate, then there is an inherent risk of reducing the amount of correct information provided as well (Koriat & Goldsmith, 1996). The results, however, allay these concerns. In both the biased and unbiased lineup instructions conditions, the availability of the not sure option came at no significant cost to the target identification rate, indicating that the not sure option does not impair memory performance. Koriat and Goldsmith's metamemory model predicts that raising the decision criterion should reduce the number of correct identifications. Indeed, Steblay & Philips (2011) in a study examining the influence of the not sure option on eyewitness decisions in simultaneous and sequential lineups found a decrease in choosing rates across correct identifications and filler identifications. However, this effect was found only for sequential lineups, and not for simultaneous lineups. Other recent studies on the influence of the *not sure* option on

memory accuracy have on the other hand produced results showing little to no significant costs to the target identification rate (e.g., Koriat & Goldsmith, 1996; Brewer, Keast & Sauer, 2010; Weber & Perfect, 2011). By showing that the *not sure* option does not lead to a decrease in target identifications in a simultaneous lineup, the present study's results add to the growing chorus of recent evidence showing that allowing witnesses the option to withhold a response can improve memory accuracy, without necessarily leading to a reduction in the amount of correct information provided. The addition of the *not sure* option had virtually no effect on the target identification rate, thereby implying that the *not sure* option went beyond providing witnesses a means of simply hesitating to make a response.

The results suggest that the *not sure* option may do more than simply raise the decision criterion of witnesses. If the sole purpose of the *not sure* option was to raise the decision criterion, there should be a corresponding drop off in hits in target present lineups. This is predicted by Koriat and Goldsmith's metamemory reporting model, which assumes that people use a singular decision criterion in making decisions. The present study's pattern of results can be explained through an alternative model: Clark's (2003) WITNESS model. The model states that there are two decision criteria; one for acceptance, and one for rejection. An identification is made if a response exceeds the criterion for acceptance, and a rejection is made if a response falls below the criterion for rejection. The *not sure* option will have the effect of increasing the criterion for making a rejection. The resulting effect is that the *not sure* response is used by witnesses who would otherwise have made both false identifications *or* false rejections.

Previous studies investigating the influence of a not sure option in lineup tasks have focused to a large extent on hits, false identifications, and the diagnosticity estimates of hits and not sure responses (e.g., Weber & Perfect, 2011), but have underexamined how filler identifications are affected by allowing witnesses the option to withhold responding. The influence of a *not sure* option on filler identifications is important to examine for two reasons: (a) there is an advantage to reducing filler identifications in order to avoid ruining the credibility of a witness if shown a subsequent lineup; and (b) it allows us to know whether people responding not sure would have identified a filler or would have made no identification or a correct identification had they not had the *not sure* option. The results do show a significant reduction in filler identifications in both target absent and target present lineups with the addition of a not *sure* option. The reduction in false identifications, coupled with a corresponding reduction in filler identifications when a not sure option is available shows that witnesses who would have otherwise made inaccurate responses (filler identifications) opt for a *not* sure response when it is available. Showing that the availability of a not sure option reduces filler identifications provide additional support for the practical benefits of allowing witnesses the option to withhold reporting a response. For instance, witnesses may be shown multiple lineups, and identifying a filler as the culprit from the first lineup may 'ruin' the witness for subsequent lineups. Even if the correct identification rate and false identification rate were only marginally affected by the provision of a *not sure* option, reducing the filler identification rate is a significant often-overlooked benefit of a given lineup procedure.

Another line of inquiry was whether the *not sure* option would have a similar effect when lineup instructions are biased versus unbiased. For instance, unbiased instructions raise the decision criterion, thereby possibly rendering the *not sure* option redundant. Therefore, it is useful to examine the extent to which the *not sure* option influences the false identification rate in spite of the potential criterion-altering effects of unbiased lineup instructions. The results show that the addition of a *not sure* option led to an approximately fifty percent reduction in choosing rate and false identification rate regardless of whether the instructions were otherwise biased or not, clearly showing that the *not sure* option is not redundant with unbiased lineup instructions. This presents a case for including an explicit *not sure* option in addition to existing lineup procedures that already use unbiased instructions.

The results showing that the provision of the *not sure* option leads to a reduction in false identifications, without a corresponding reduction in target identifications, is significant for its implication on diagnosticity. Diagnosticity is determined by examining the proportion of hits relative to the proportion of false identifications. Accordingly, diagnosticity of an identification response should increase to the extent that the *not sure* option maintains the hit rate, while decreasing the false ID rate. While the diagnosticity of an identification increased in absolute value with the addition of a *not sure* option (e.g., from 7.52 to 15.78 when lineup instructions were unbiased), this difference was not significant. However, finding differences in diagnosticity requires a large sample size; this non-significant difference is likely reflective of a lack of power. Indeed, looking at hits and false alarms separately did show a significant advantage of the *not sure* option. Future research using a larger sample size is needed to further examine this issue.

Further information gain analyses revealed that *not sure* responses do provide significant information about the likelihood of guilt or innocence of a suspect. The information gain curve for *not sure* responses significantly differed from zero, and provided exonerating evidence. This mirrors Wells and Olson's (2002) findings that witnesses were more likely to respond *not sure* when the perpetrator was not in the lineup, than when the perpetrator was in the lineup. Therefore, in addition to its effects on witness accuracy, a *not sure* response enables the examination of the likelihood of innocence of a suspect.

Koriat and Goldsmith's (1996) metamemory reporting model provided a framework within which memory decisions on tasks can be examined. In their original study, they found that allowing witnesses the option to withhold a response produced higher levels of accuracy on a general knowledge task. Using a lineup task, the present study found similar beneficial effects of the *not sure* option on witness accuracy, leading to the conclusion that these benefits may be generalizable across different tasks. In the eyewitness literature, different studies have shown that the *not sure* option can reduce false identifications without reducing hits for participants who viewed show-ups and simultaneous lineups. However, one recent study directly examining the influence of the *not sure* option found a reduction in hits when participants viewed sequential lineups (Steblay & Philips, 2011). This raises questions as to whether the beneficial effect of the *not sure* option is thus limited to only show-ups and simultaneous lineups. Future research is needed to examine the extent to which these beneficial effects can be generalized across the various procedures.

Overall, the results conform well with data from the recent studies examining the influence of a *not sure* option on accuracy. Weber and Perfect's (2011) study revealed significant benefits for the provision of a *not sure* option. Similar to the present study, their study found that the *not sure* option reduced false identifications without a concomitant reduction in correct identifications. A key procedural difference is that they used a showup, whereas the present study used a simultaneous lineup. The present study extends their findings by virtue of the fact that similar results were found using a simultaneous lineup, suggesting that the beneficial effects of the *not sure* option may be generalizable across different identification procedures.

Using lineups, Steblay and Philips' (2011) study revealed mixed results. While the *not sure* option was able to reduce false identifications, it also reduced correct identifications. Importantly, this effect was found only for witnesses who viewed sequential lineups. While the availability of the *not sure* option did not lead to a significant reduction in correct identifications for those who viewed a simultaneous lineup, it did not lead to a reduction in false identifications either. Similarly, Brewer, Keast and Sauer's (2010) study showed only marginal benefits for the provision of a *not sure* option in a simultaneous lineup. A possible explanation for the discrepancy in results between the two studies and the present study is that the present study used a stronger manipulation of the *not sure* option. In the present study, the *not sure* option was manipulated both as an instruction displayed on the witness' computer screen prior to viewing the lineup, and as a response option available to witnesses. Earlier studies only presented it as a response option, without also providing it in a pre-lineup instruction.

Additionally, information gain analyses from the present study show results that are consistent with Wells & Olson's (2002) findings. As with their data, information gain for *not sure* responses was significantly greater than zero, and tended to provide exonerating evidence. In summary, the present study bolsters evidence from the aforementioned recent studies and expands on their findings showing advantages for providing witnesses with a *not sure* response option.

It is plausible that the influence of the *not sure* option may be moderated by memory strength. For instance, when memory strength is strong, it is possible that raising the decision criterion would have less influence, compared to when memory strength is weak. Consequently, *not sure* responses may be rendered to a greater degree when memory is weak, rather than strong. Future research on this topic should examine the role of memory strength on the impact of a *not sure* option. Future researchers can for instance examine the impact of memory strength in varying ways, such as by manipulating witness' exposure time to the face of the perpetrator.

Overall, the results suggest that the *not sure* option had a direct and significant impact on eyewitness decisions. An explicit *not sure* option, as evidenced by the results, provides enormous benefits by reducing false identifications and filler identifications, thereby improving overall accuracy of witnesses. In addition to improving witness accuracy, a *not sure* response in and of itself provides significant information that is of exonerating value. Showing that the *not sure* option provides significant informational value further bolsters the case for integrating this procedure in police lineup administration.

Hitherto, few studies in the eyewitness literature show manipulations that successfully reduce false identifications and filler identifications, without reducing hits. The present study provides evidence of how a simple-to-implement manipulation can achieve these aims. From a practical perspective, providing an option for witnesses to opt out of making likely incorrect identifications does not necessarily reduce the amount of correct identifications. Results from the present study have shown that the *not sure* option does not impair identification accuracy. In fact, it improves it. Given the paucity of research on this issue, further research is needed to assess the generalizability of these findings, and the conditions under which these effects can be maximized. Nonetheless, allowing witnesses the option to withhold making an identification response can reduce errors of potentially grave consequences.

REFERENCES

- Brewer, N., Keast, A., & Sauer, J. (2010). Children's eyewitness identification performance: Effects of Not Sure and accuracy motivation instructions. *Legal and Criminological Psychology*, 15, 261-27.
- Charman, S. D., & Cahill, B. S. (2012). Witnesses' memories for lineup fillers postdict their identification accuracy. *Journal of Applied Research in Memory and Cognition*, *1*, 11–17.
- Charman, S. D., & Wells, G. L. (2007). Eyewitness lineups: Is the appearance-change instruction a good idea? *Law and Human Behavior*, *31*, *3–22*.
- Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology*, 17, 629–654.
- Clark, S.E., Howell, R., & Davey, S.D. (2008). Regularities of eyewitness identification. *Law and Human Behavior*, 32, 406-422.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, 29, 575–604.
- Cutler, B. and Penrod, S. (1995) *Mistaken Identity*. Cambridge University Press: Cambridge UK.
- Ebbesen, E.B., & Flowe H.D. (2002). Simultaneous v. sequential lineups: what do we really know? Retreived 20 November 2011, from http://www-psy.ucsd.edu/%7eeebbesen/SimSeq.htm.
- Goldstein, Herman. 1977. Categorizing and structuring discretion. Chapter 5 (pps. 93-130) in Policing a Free Society. Cambridge, MA: Ballinger Publishing Company.
- Kohnken, G., & Maass, A. (1988). Eyewitness testimony: False alarms on biased instructions? *Journal of Applied Psychology*, 73, 363-370.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103, 490-517.
- Loftus, E. F. (1979). Eyewitness testimony. Cambridge, MA: Harvard University Press.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, 51, 102–116.
- Pozzulo, J. D. & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: Ameta-analysis. *Law & Human Behavior*, 22, 549-570.

- Rhodes, M. G., & Kelley, C. M. (2005). Executive processes, memory accuracy, and memory monitoring: An aging and individual differences analysis. *Journal of Memory and Language*, 52, 578-594.
- Ross M, Buehler R. 1994. Creative remembering. In The Remembering Self, ed. U Neisser, R Fivush, pp. 205–35. New York: Cambridge Univ. Press.
- Shaw, J. S., III. (1996). Increases in eyewitness confidence resulting from post event questioning. *Journal of Experimental Psychology: Applied*, 2, 126-146.
- Spehn, M.K. & Reder, L.M. (2000). The unconscious feeling of knowing: A commentary on Koriat's paper. *Consciousness and Cognition*, 9, 187-192.
- Steblay, N.M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior*, 21, 283-297.
- Steblay, N., Dysart, J., Fulero, S. & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, 25, 459-473.
- Weber, N., & Perfect, T. J. (2011, February 17). Improving Eyewitness Identification Accuracy by Screening Out Those Who Say They Don't Know. *Law and Human Behavior*. Advance online publication. doi: 10.1037/h009397
- Wells, G.L., & Bradfield, A.L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology*, 83, 360 - 376.
- Wells, G. L. & Loftus, E. F. (2003). Eyewitness memory for people and events. In A. Goldstein, Ed. Comprehensive handbook of psychology, Volume 11, Forensic psychology. New York: John Wiley and Sons.
- Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, 88, 776–784.
- Wells, G. L., & Olson, E. A. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied*, 8, 155–167.
- Wells, G. L., & Olson, E. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54, 277-295.
- Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, 33, 1–24.

Lineup Type		Lineup Instru	Lineup Instruction			
		Biased	Unbiased	Overall		
ТР						
	п	4	3	7		
	cell n	30	33	63		
	%	13.3	9.1	11.1		
TA						
	п	20	9	29		
	cell n	34	32	66		
	%	58.8	28.1	43.9		

Table 1 Frequency and percentage of *not sure* responses, and cell n, by lineup type and lineup instruction

Lineup Instruction		Not Sure Option			
		Yes	No	Overall	
Biased					
	n	18	17	35	
	cell n	30	31	61	
	%	60	54.8	57.4	
Unbiased					
	n	19	17	36	
	cell n	33	30	63	
	%	57.6	56.6	57.1	

Table 2 Percentage of correct identifications in TP lineups, and cell n, by not sure option

Lineup Instruction	Not Sure Option		
	Yes	No	Overall
Biased			
п	14	30	44
cell n	34	30	64
%	41.2	100	68.7
Unbiased			
n	7	14	21
cell n	32	30	62
%	21.9	45.2	33.9

Table 3 Percentage of false identifications in TA lineups, and cell n, by not sure option



Figure 1 Information gain from *not sure* responses for biased versus unbiased lineup instructions



Figure 2 Participants who received biased lineup instructions: Information gain from identification of suspect for when *not sure* option is available versus when *not sure* option is not available



Figure 3 Participants who received unbiased lineup instructions: Information gain from identification of suspect for when *not sure* option is available versus when *not sure* option is not available

	Not Sure	Standard (no not sure)
Biased Instructions	<i>"From the lineup presented, identify the perpetrator. You are permitted to pick him out, or you may answer 'not sure'"</i>	" From the lineup presented, identify the perpetrator. You are permitted to pick him out, or answer that he is not there"
Unbiased Instructions	"The lineup you are about to view may or may not contain the perpetrator. You are permitted to pick him out if he is there, answer that he is not there, or you may answer 'not sure'"	"The lineup you are about to view may or may not contain the perpetrator. You are permitted to pick him out if he is there, or answer that he is not there"

Appendix A. *Instructions associated with the conditions in TA and TP lineups.*