2017

Eyewitness identification performance on showups improves with an additional-opportunities instruction: Evidence for present–absent criteria discrepancy

Andrew M. Smith
Carleton University

Gary L. Wells
Iowa State University, glwells@iastate.edu

R. C. L. Lindsay
Queen’s University

Tiffany Myerson
York University

Follow this and additional works at: https://lib.dr.iastate.edu/psychology_pubs

Part of the Applied Behavior Analysis Commons, Criminology and Criminal Justice Commons, Law and Psychology Commons, and the Theory and Philosophy Commons

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/psychology_pubs/63. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Eyewitness identification performance on showups improves with an additional-opportunities instruction: Evidence for present–absent criteria discrepancy

Abstract
We tested the proposition that when eyewitnesses find it difficult to recognize a suspect (as in a culprit-absent showup), eyewitnesses accept a weaker match to memory for making an identification. We tie this proposition to the basic recognition memory literature, which shows people use lower decision criteria when recognition is made difficult so as to not miss their chance of getting a hit on the target. We randomly assigned participant–witnesses (N = 610) to a condition in which they were told that if they did not believe the suspect was the culprit, they would have additional opportunities to make an identification later (additional-opportunities instruction). We fully crossed this instruction with the standard admonition (i.e., the culprit may or may not be present) and with the presence or absence of the culprit in a showup identification procedure. The standard admonition had no impact on eyewitness decision-making; however, the additional-opportunities instruction reduced innocent-suspect identifications (from 33% to 15%) to a greater extent than culprit identifications (57% to 51%). The additional-opportunities instruction yielded a better tradeoff between culprit and innocent-suspect identifications as indicated by binary logistic regression and receiver operator characteristic (ROC) analyses.

Keywords
Eyewitness Decision-Making, Present/Absent Criteria Discrepancy, Showups, Signal Detection Theory, Decision Criterion

Disciplines
Applied Behavior Analysis | Criminology and Criminal Justice | Law and Psychology | Psychology | Theory and Philosophy

Comments
©American Psychological Association, 2018. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at doi:10.1037/lhb0000284. Posted with permission.
Eyewitness Identification Performance on Showups Improves with an Additional-Opportunities Instruction: Evidence for Present/Absent Criteria Discrepancy

Andrew M. Smith  Gary L. Wells  R. C. L. Lindsay  Tiffany Myerson

This article has been accepted for publication in *Law and Human Behavior*.

© 2018. American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission.

Author Note

Andrew M. Smith, Department of Psychology, Carleton University; Gary L. Wells, Department of Psychology, Iowa State University; R. C. L. Lindsay, Department of Psychology, Queen’s University; Tiffany Myerson, Osgoode Hall Law School, York University.

Correspondence concerning this article should be addressed to Andrew M. Smith, Department of Psychology, Carleton University, 550 Loeb Building, 1125 Colonel By Drive, Ottawa, ON, Canada K1S 5B6. AndrewMSmith@cunet.carleton.ca

The authors would like to acknowledge Norwood Police Department, Chief of Police, William G. Brooks for discussing his showup protocol with us.
Abstract

We tested the proposition that when eyewitnesses find it difficult to recognize a suspect (as in a culprit-absent showup), eyewitnesses accept a weaker match-to-memory for making an identification. We tie this proposition to the basic recognition memory literature, which shows people use lower decision criteria when recognition is made difficult so as to not miss their chance of getting a hit on the target. We randomly assigned participant-witnesses ($N = 610$) to a condition in which they were told that if they did not believe the suspect was the culprit, they would have additional opportunities to make an identification later (additional-opportunities instruction). We fully crossed this instruction with the standard admonition (i.e., the culprit may-or-may-not be present) and with the presence or absence of the culprit in a showup identification procedure. The standard admonition had no impact on eyewitness decision-making; however, the additional-opportunities instruction reduced innocent-suspect identifications (from 33% to 15%) to a greater extent than culprit identifications (57% to 51%). The additional-opportunities instruction yielded a better tradeoff between culprit and innocent-suspect identifications as indicated by binary logistic regression and Receiver Operator Characteristic (ROC) analyses.

Keywords: Eyewitness Decision-Making; Present/Absent Criteria Discrepancy; Showups; Signal Detection Theory; Decision Criterion
Public Significance Statement: We told some eyewitness-participants that if they did not believe the suspect in a show-up was the culprit, they would have additional opportunities to identify someone later. This instruction decreased mistaken identifications more than it reduced accurate identifications. This finding and its theoretical underpinnings shows new promise for pre-identification instructions that can increase the reliability of eyewitness identification evidence.
Eyewitness Identification Performance on Showups Improves with an Additional-Opportunities Instruction: Evidence for Present/Absent Criteria Discrepancy

Based on reviews of data on pre-lineup instructions, it has been argued that an instruction admonition (that the culprit might not be in the lineup) reduces mistaken identifications at a greater rate than it reduces culprit identifications, thereby resulting in an overall improvement in performance (Steblay, 1997, 2013). Clark (2005), however, argued that finding an improvement in performance from instructions could not be reconciled with Signal Detection Theory (Green & Swets, 1966; Macmillan & Creelman, 2005). Clark (2005) argued that instructions could only produce an equal tradeoff in culprit and innocent suspect identifications. Clark’s (2005) logic is that instructions influence the tendency for an eyewitness to make an identification (response bias), but not the ability of an eyewitness to distinguish between guilty and innocent persons (discriminability). Because instructions should only influence response bias, Clark (2005) reasoned that, when measured by statistics that are independent of response bias (e.g., the discriminability index or $d'$) instructions could only produce a tradeoff between culprit and innocent-suspect identifications.

We propose, however, that an instruction can improve performance as measured by $d'$ (or the area under the ROC curve, AUC) in an eyewitness identification task. In the present experiment, we test this proposition with a novel instruction admonishing eyewitnesses that if they do not believe the suspect is the culprit they will have additional opportunities to make an identification later. We fully crossed this instruction with the presence and absence of the culprit in a showup identification procedure and with a standard admonition informing the eyewitness that the culprit may or may not be present.
Hereafter, we will argue that eyewitnesses who encounter innocent suspects tend to require a weaker match-to-memory for positive identification (i.e., adopt a lower decision criterion) than do eyewitnesses who encounter guilty suspects. We refer to this as present/absent criteria discrepancy theory. To be clear from the outset, it is not our contention that eyewitnesses knowingly lower their criteria for identification because they are aware that the suspect is innocent. Rather, when eyewitnesses lower their decision criteria, they do so because they find the recognition task difficult or because match-to-memory is weak. Because innocent suspects tend to provide a worse match-to-memory than do guilty suspects, eyewitnesses who encounter innocent suspects might have a tendency to find the recognition task more difficult and to lower their criteria in response to this increased task difficulty.

Although it is generally the case that instructions affect only decision criteria and not discriminability, eyewitness identification tasks have a property that does not characterize any of the extant literature and methods typically used with Signal Detection Theory. Specifically, perception and memory studies routinely present each participant with a large number of trials (dozens or sometimes hundreds) in which the signal is present on some trials and absent on other trials. In other words, the routine method in the basic recognition and perception literatures involves using a repeated-measures task in which each participant is exposed to a randomly-ordered sequence of signal-present and signal-absent test trials. This is very unlike an eyewitness identification experiment in which the participant is tested with a single trial in which the target-present versus target-absent factor is manipulated between participants. In other words, in a standard eyewitness identification study, each eyewitness is nested completely in either a signal-
present trial or a signal-absent trial and never experiences the other state of the world. Other than eyewitness identification experiments, we can think of no other type of cognitive psychology experiment that has the property of nesting the participant in a single-trial task in which the signal is either present or absent.

This distinction between a basic perception/memory experiment using within-participant, repeated-measure designs, over a large number of signal-present and signal-absent trials versus an eyewitness identification experiment in which the participant is tested once using either a signal-present or a signal-absent trial is potentially very important. In fact, in standard signal detection experiments, discriminability and response bias are calculated individually for each participant and these values become the unit of analysis for inferential comparison of performance under different conditions. In an eyewitness identification experiment, however, discriminability cannot be calculated at the level of the individual participant because any given participant has only one data point and those in the signal-present condition never experience the signal-absent condition and vice versa.

The fact that the eyewitness experiences only a single test trial of signal present or signal absent raises an important question. Specifically, in an eyewitness identification experiment, can we assume that the decision criterion set by an eyewitness who encounters a target-present (signal) test trial is the same as the decision criterion set by an eyewitness who encounters a target-absent (noise only) test trial? The presumption of Signal Detection Theory is that the decision criterion that a decision-maker uses for signal-absent test trials is the same as it is for signal-present test trials (e.g., Green & Swets, 1966). In the case of a showup (a one-person lineup), for example, the
premption would be that an eyewitness would use the same criterion (i.e., how similar to my memory does the suspect need to be for me to identify the person?) if she were shown a target-present showup as if she were shown a target-absent showup. In a repeated-measures experiment (which often uses initial practice trials) it makes sense that a person would settle on a criterion after a few test trials and stay with that criterion across remaining trials whether the signal is present or absent. But, we argue, in a single-trial test we cannot assume that those who were given a target-absent test used the same decision criterion as those who were given a target-present test.

In fact, there are various forms of evidence in basic cognitive psychology that suggest that eyewitnesses who encounter a target-absent test will use a lower criterion than will eyewitnesses who encounter a target-present test. In word-memory experiments, for example, participants will lower their decision criterion when they are tested on word blocks for which they have difficulty selecting a signal due to having had fewer study trials on the words used in that test block (e.g., Verde & Rotello, 2007). Similarly, researchers have noted that the so-called strength-based mirror effect, in which conditions that weaken participants’ memory experience lead to not only fewer hits but also more false alarms, is due to people responding to the difficulty of detecting signal by lowering their decision criterion in order to avoid missing a hit (Wixted & Gaitan, 2002). Hence, as the memory task becomes more difficult (weaker sense of a recognition signal), people lower their decision criterion. Lowering decision criteria when the memory task becomes more difficult is not just true for word memory experiments but also for eyewitness identification experiments. For example, exposure duration (Memon, Hope, & Bull, 2003), distance at encoding (Lampinen, Erickson, Moore, & Hittson, 2014), and disguise
PRESENT/ABSENT CRITERIA DISCREPANCY

(Mansour, Beaudry, Bertrand, Kalmet, Melsom, & Lindsay, 2012) affect eyewitnesses’ decision criteria. In each case, the condition associated with less optimal encoding conditions (shorter exposure, greater distance at encoding, a disguised target) results in eyewitnesses using lower decision criteria. This lowering of decision criterion in the face of difficult recognition tasks occurs not only when the difficulty manipulation occurs at encoding conditions but also when the difficulty manipulation occurs at the time of the recognition test (Hintzman, Caulton, & Curran, 1994).

Clearly, when an eyewitness encounters a target-absent identification procedure (innocent suspect) the eyewitness will experience more difficulty detecting the signal than when the eyewitness encounters a target-present (guilty suspect) procedure. Indeed, in the target-absent procedure, there is no signal to detect. Yet, if eyewitnesses assume that this is their only opportunity to identify the culprit, difficulty in a culprit-absent condition might lead them to lower their criterion in order to make sure they have a chance to make a hit. Of course, eyewitnesses do not drop their criterion because they think they are in a culprit-absent condition. All the eyewitnesses know is that they are having a weak recognition experience and it is their task to determine the cause of the weak recognition experience. To the extent that eyewitnesses misattribute the weak recognition experience to something other than the absence of the culprit (e.g., to a poor view or to having paid little attention), eyewitnesses might drop their criterion and mistakenly identify an innocent person. In fact, research shows that after viewing a simulated crime, witnesses who view a randomly-assigned culprit-absent lineup report having had a worse view and having paid less attention than do witnesses who view a randomly-assigned culprit-present lineup (Bradfield, Wells, & Olson, 2002). In other
words, culprit-absent witnesses tend to misattribute their difficulty in detecting the culprit in the lineup to having had a poor witnessing view and having not paid much attention at the time of witnessing rather than attributing their difficulty to the fact that they had been randomly assigned to a culprit-absent lineup. But these metacognitions were incorrect – those randomly assigned to culprit-present versus culprit-absent lineups had the same view and on average paid the same amount of attention. The work of Palmer, Brewer, and Weber (2010) has shown the important role that these metacognitions can play in the processes by which eyewitnesses make their identification decisions.

Our theory posits that there is a propensity for eyewitnesses to set a lower decision criterion when they encounter a culprit-absent procedure than when they encounter a culprit-present procedure just as people in other tasks set a lower decision criterion when the signal is difficult to detect due to weak memories. In both cases, the reason they do this is because they fear missing the opportunity to get a hit. Indeed, when signal strength weakens it is necessary to lower one’s decision criterion in order to avoid too many misses. But if this weak signal and task difficulty is due to the fact that the culprit is absent from the identification procedure, then lowering the decision criterion does not increase hits and instead purely increases false alarms. This setting of a lower decision criterion for a target-absent identification procedure than for a target-present identification procedure is what we call present/absent criteria discrepancy.

Importantly, present-absent criteria discrepancy has the effect of profoundly undermining memory performance. Whether the decision criterion is high or low, applying the same criterion to both the absent and present conditions will produce better performance than will setting a lower criterion for absent conditions than for present
Present/Absent Criteria Discrepancy

Conditions. Setting a lower decision criterion for culprit-absent conditions than for
culprit-present conditions is extremely damaging to overall performance. Consider the
probability of occurrence curves in Panels A and B of Figure 1, for example. The black
curves reflect the range of match-to-memory values that eyewitnesses who encounter the
culprit might have and the grey curves reflect the range of match-to-memory values that
eyewitnesses who encounter innocent suspects might have. Because the culprit will tend
to provide a better match-to-memory than will an innocent suspect, the black distribution
exceeds the grey distribution. Discriminability is reflected by the degree of overlap
between the culprit and innocent distributions and is identical in Panels A and B. In Panel
A, the single black vertical line reflects the decision criterion. If the suspect’s match
value exceeds this criterion, the eyewitness makes an ID. If not, the eyewitness rejects.
But, if eyewitnesses who encounter culprit-absent procedures tend to use more lenient
criteria than eyewitnesses who encounter culprit-present procedures, then we need a
second, lower criterion, for eyewitnesses in the culprit-absent condition. The grey vertical
line in Panel B reflects this lower criterion. The result of using a lower criterion in
culprit-absent than in culprit-present conditions is an increase in false positives without a
concomitant increase in hits and, as a result, much poorer performance overall.

This brings us back to the debate about whether instructions can improve
discriminability as measured by signal-detection statistics such as the AUC or $d’$. We
agree with Clark (2005) that if the decision criterion is the same in culprit-present and
culprit-absent procedures, instructions should not improve discriminability (the distance
between the innocent and guilty suspects distributions in Figure 1) and instead should
only affect criterion placement. But, what if witnesses tend to set a lower criterion when
the culprit is absent than when the culprit is present and there was an instruction that mitigated the propensity of witnesses to set a lower criterion when the culprit is absent? If there were an instruction that lessened the propensity to set a lower criterion in the absent than in the present conditions, then the tradeoff between culprit and innocent suspect identifications as measured by the AUC or $d'$ could be improved without changing the distance between innocent and guilty suspect distributions.

The key to fashioning such an instruction requires an understanding of why witnesses set a lower criterion for culprit-absent identification procedures in the first place. We noted earlier that when witnesses encounter a culprit-absent identification procedure, they experience a weak match-to-memory sensation. Some witnesses will accurately attribute this to the fact that the culprit is absent, but many will misattribute the weak sense of match to other things, such as their attention and view having been weak (see Bradfield et al., 2002). Accordingly, these witnesses will lower their decision criterion in order to not miss what they believe to be their only chance to identify the culprit. Our account parallels the accounts that others have used to explain why people in word-memory experiments lower their criterion when their memories are weak (e.g., Morrell, Gaitan, & Wixted, 2002; Wixted & Gaitan, 2002).

In our attempt to devise an instruction that would mitigate the tendency of witnesses to lower their criterion in the absent condition while having little effect on criterion setting in the present condition, we told half of our witnesses that if they did not think the culprit was in the showup, they would have an additional opportunity to view someone else later. If witnesses in culprit-absent conditions are setting a low criterion because they do not want to miss the opportunity to identify the culprit, this instruction
should mitigate that tendency. After all, the additional-opportunities instruction helps witnesses realize that lowering their criteria so as to choose a person in the first procedure might actually prevent them from hitting on the actual culprit in a later identification procedure. Hence, this instruction should reduce mistaken identifications in culprit-absent conditions. At the same time, however, this additional-opportunities instruction should have less impact on culprit identifications. Indeed, the instruction is targeted at eyewitnesses who are having a relatively weak match-to-memory experience, but choose to make an identification out of concern that this is their only opportunity to identify the culprit. On average, eyewitnesses who are viewing a culprit should have a stronger match-to-memory experience and so fewer culprit identifications than innocent-suspect identifications should be attributable to the eyewitness’ fear that this is the only opportunity to identify the culprit.

**Method**

The Institutional Review Board for human research at Iowa State University approved this experiment.

**Participants and Design.** We recruited undergraduates \((N = 610)\) from the Psychology Participant Pool at Iowa State University in exchange for course credit. Participants, on average, were 19.23 (SD = 1.82) years of age and 64% were female. The majority of participants were of European ancestry: European (79.34%), Hispanic (3.77%), Asian or Pacific Islander (9.18%), Black (2.46%), and “Other” (5.25%).

Participants were randomly assigned to a 2 (may-or-may-not admonition: no, yes) \(\times\) 2 (additional-opportunities instruction: no, yes) \(\times\) 2 (Showup: target absent, target present) between-subjects design. We also examined performance in iterative-showup
procedures. Replicating Smith et al. (2014), we found that cautionary instructions reduced innocent suspect identifications for iterative-showup procedures. But, all of the action from the additional-opportunities instruction occurred on the first showup, so we do not discuss iterative-showups further. Interested readers are referred to the supplementary materials.

Materials.

**Targets.** Each participant was presented with only a single target video; however, we used three different targets in this study to attain some limited degree of stimulus sampling (Wells & Windschitl, 1999). Each target was yoked to three innocent suspects who each fit the general description of their respective target. All targets and respective innocents were male, European, in their early 20s, and varied in hair length and color (very short, light brown hair; medium-length, very dark hair; medium-length blonde hair). None of our outcome measures varied as a function of stimuli; therefore, we do not differentiate among them in our analyses.

Each target video lasted approximately 2s in duration, depicted the target’s face from a three-quarter angle, and included a brief statement from the target related to criminal activity (e.g., “So, there’s good stuff in that place?”). Each video was approximately 18 cm × 25 cm. In each showup photo, the face of the suspect was depicted from the neck up (no clothing cues) looking directly at the camera. Each photograph was approximately 22 cm x 28 cm. The choosing pattern did not vary as a function of target stimuli, so we do not make any further distinctions among these stimuli.
**Filler task.** After the encoding task, participants provided informed consent, completed a demographics questionnaire, completed the self-monitoring scale (Snyder, 1974) for an unrelated study, and completed a visual search task involving a beach scene from the children’s book, “Where’s Waldo?” ¹ Participants were given a sheet of paper with a list of questions regarding the image, such as “How many people in the picture are in red swimsuits?” The visual search task lasted eight minutes and in total, there was an approximately 10-minute delay between the encoding video and the initial show-up procedure.

**Procedure.** Participants were brought into a room, individually, and instructed to watch a video on a computer. Following the video, participants read a letter of information on screen, gave consent, and provided demographic information. After providing demographic information, participants completed the self-monitoring scale (Snyder, 1974) and completed the visual search task. The delay between encoding and recognition was approximately 10 minutes.

Participants who received neither the standard admonition nor the additional-opportunities instruction were told: “In a moment you will view a photo of a man. Please indicate whether the man in the photo is the man from the video you saw at the beginning of the experimental session.” Participants received the standard admonition were also instructed that the photo may or may not be of the man from the video. Participants who received the additional-opportunities instruction were also told that “If you do not believe that the man in the photo is the man you saw, others will be shown later for you to attempt to identify the man from the video.” After their identification decisions,

¹ TM & © 2008 Entertainment Rights Distribution Limited. All rights reserved.
participants were asked to indicate their level of confidence from 0% - 100% (in 10-point increments).

**Results**

In our primary analysis we regressed choosing on the additional-opportunities instruction, the standard admonition, and target presence. We followed up significant interactions by examining simple-main effects within target-present and -absent conditions. Significant interactions indicate a better tradeoff between culprit and innocent suspect identifications and main effects indicate more conservative responding. Keeping with traditional practice in eyewitness research, we compute likelihood ratios to assess the diagnosticity of identification decisions (Wells & Lindsay, 1980). Because some researchers have recently advocated for the exclusive use of ROC analysis to compare identification procedures, we also constructed ROC curves (e.g., Wixted & Mickes, 2012). We conclude by examining the predictive value of expressed level of eyewitness confidence with both a model-building approach and calibration analysis. ROC analyses were generated using the pROC statistical package (Robin et al., 2011) and all statistical analyses were based on the nonparametric bootstrapping method (n = 2000).

For some of our key descriptive statistics, we were only able to obtain a point estimate for each condition and had no ability to assess the variance around these statistics (e.g., the diagnosticity ratio, calibration statistics). For the diagnosticity ratio we used a Monte Carlo simulation to calculate the probability of observing a difference of equal or greater magnitude if the null were true. For the calibration statistics, we use the modified-jackknife procedure described by Horry, Palmer, and Brewer (2012). We used the jackknife procedure to estimate the standard error for each statistic (Mosteller &
Tukey, 1968) and then we used these estimated standard errors to calculate inferential 95% confidence intervals (CIs) around each statistic (Tryon, 2001). Non-overlapping inferential CIs indicate significance at the $\alpha = .05$ level.

Comparing Culprit and Innocent Suspect Identifications with Logistic Regression Analysis. The three-way interaction between the additional-opportunities instruction, the standard admonition, and target-presence was not significant, $B = -.13$, $SE = .74$, Wald’s $\chi^2(1) = 0.03$, $p = .86$, $e^B = 0.88$, (95% CI [0.21, 3.73]). Neither the two-way interaction between the standard admonition and target-presence, $B = -.13$, $SE = .36$, Wald’s $\chi^2(1) = 0.12$, $p = .73$, $e^B = 0.88$, (95% CI [0.43, 1.80]), nor the main effect of the standard admonition, $B = -.11$, $SE = .18$, Wald’s $\chi^2(1) = 0.39$, $p = .53$, $e^B = 0.90$, (95% CI [0.64, 1.27]) were significant. The standard admonition neither impacted the tradeoff between culprit and innocent suspect identifications nor the frequency with which eyewitnesses identified suspects. Accordingly, we do not consider the standard admonition further.

In contrast to the standard admonition, the additional-opportunities instruction decreased the frequency with which eyewitnesses identified suspects, $B = -.57$, $SE = .18$, Wald’s $\chi^2(1) = 10.47$, $p = .001$, $e^B = 0.56$, (95% CI [0.40, 0.80]). Participants who were provided with the additional-opportunities instruction (0.49:1) were 44% ($0.56 - 1 = -44\%$) less likely to identify the suspect than were those participants who did not receive the AOI (0.83:1). In other words, the additional-opportunities instruction produced a conservative shift in response bias. Importantly, this main effect was qualified by a significant interaction between the additional-opportunities instruction and target-presence, $B = .85$, $SE = .37$, Wald’s $\chi^2(1) = 5.27$, $p = .02$, $e^B = 2.33$, (95% CI [1.13,
The additional-opportunities instruction significantly decreased innocent suspect identifications, $B = -1.08$, $SE = .29$, Wald’s $\chi^2(1) = 14.00$, $p < .001$, $e^B = 0.34$, (95% CI [0.19, 0.60]), but did not significantly decrease culprit identifications, $B = -.23$, $SE = .23$, Wald’s $\chi^2(1) = 1.01$, $p = .32$, $e^B = 0.79$, (95% CI [.51, 1.25]). Participants who received the AOI ($P = .15$, Odds = 0.17:1) were 66% ($0.34 – 1 = -66\%$) less likely to identify the innocent suspect than were those participants who did not receive the additional-opportunities instruction ($P = .33$, Odds = 0.50:1). Hence, the additional-opportunities instruction produced a better tradeoff between culprit and innocent-suspect identifications. All proportions are provided in Table 1.

**The Diagnosticity Ratio.** Next we examined the impact of the additional-opportunities instruction on the diagnosticity ratio. Because the underlying distributions for ratio-based measures inherently suffer from positive skew, we log transformed the respective diagnosticity ratios and inferentially compared the log-likelihoods. This is analogous to the process employed when using logistic regression analysis, for which the log-odds are subjected to the inferential test before being exponentiated back into an odds ratio to make interpretation more intuitive. Suspect identifications made in the presence of the additional-opportunities instruction were more diagnostic (diagnosticity ratio = 3.52) than were suspect identifications made in the absence of the additional-opportunities instruction (diagnosticity ratio = 1.71), $p < .001$.

**ROC Analysis.** We next examined the additional-opportunities instruction with ROC analysis. For identification procedures, ROC curves are constructed by plotting culprit identifications against innocent suspect identifications at cumulating levels of confidence. The leftmost-point of the ROC curve represents culprit and innocent suspect
identifications made with 100% confidence. The second to leftmost point on the ROC curve represents all of those identifications from the first point (viz. identifications made with 100% confidence) plus all identifications made with 90% confidence. One continues plotting points in this manner until all identifications are represented in the rightmost point of the ROC curve. When comparing two ROC curves (as we will do below), the procedure that produces a higher ROC curve is the procedure that produces a better tradeoff between culprit and innocent-suspect identifications.

Because we examined showups in the present experiment, we also included eyewitnesses who made rejections in our ROC curves. This is how ROC curves are typically constructed in the basic memory literature and the result is an ROC curve that covers the entire unit square (e.g., Macmillan & Creelman, 2005). The procedure works like this: after all of the choosers (even those who made an identification with 0% confidence) are reflected in a single point on the ROC curve, the next point includes all of the choosers and the eyewitnesses who rejected the suspect with 0% confidence. At the applied level, one could interpret this as meaning that, if the eyewitness rejected the suspect, but had 0% confidence in that rejection, police might be willing to treat that as an identification instead of a rejection. One continues cumulating points in this manner until all eyewitness-participants (including those who rejected the suspect with 100% confidence) are represented in the rightmost point of the ROC curve, which will fall on the Northeast corner of the unit square (X = 1, Y = 1).

Our ROC curves are depicted in Figure 2. We first compared the ROC curves considering all eyewitness decisions, including rejections. Although the additional-opportunities instruction produced a descriptively larger AUC (.74) than the control
condition (AUC = .70), the difference was not statistically reliable, $D = 1.02$, $p = .31$. But, when ROC curves intersect (as ours do) or when there is an *a priori* reason to examine only a portion of the unit square (e.g., consider only suspect identifications), examining partial AUCs makes good sense (McClish, 1989; Zhou, O’Malley, & Mauri, 2007). We first examined a pAUC region that included all suspect identifications. Because the highest innocent suspect identification rate was .33, we examined the pAUC region of 0 - .34. In this region of the ROC space, the expected difference emerged: identifications made in the presence of the additional-opportunities instruction were more diagnostic (pAUC = .17) than identifications made in the absence of the additional-opportunities instruction (pAUC = .12), $D = 2.21$, $p = .03$. We also compared the two ROC curves in a more restricted pAUC range that extended only to the innocent-suspect identification rate in the presence of the AOI (pAUC: 0 - .15) and again, identifications made in the presence of the additional-opportunities instruction were more diagnostic (pAUC = .06) than identifications made in the absence of the additional-opportunities instruction (pAUC = .03), $D = 2.75$, $p = .006$. This is a particularly promising sign for the additional-opportunities instruction, because in this portion of the unit square, all eyewitnesses who made suspect identifications in the presence of the additional-opportunities instruction (even those who did so at the lowest levels of confidence) are being compared to only the most confident eyewitnesses from the control condition (70% - 100% confidence).

**Eyewitness Confidence.** Next, we examined how well eyewitness confidence sorted accurate and inaccurate choosers and non-choosers. See Table 2 for mean confidence as a function of choosing, accuracy, and the additional-opportunities
instruction. Table 2 also includes point-biserial correlation coefficients for the relationship between confidence and accuracy as a function of choosing and the additional-opportunities instruction.

**Confidence-Accuracy Calibration.** Calibration curves involve plotting the percent of accurate decisions at some expressed level of confidence (objective accuracy) against that expressed level of confidence (subjective accuracy) (e.g., Brewer & Wells, 2006; Cutler & Penrod, 1989; Juslin Olsson, & Winman, 1996; Wixted & Wells, 2017). Perfect calibration occurs when 100% of decisions that are made with 100% confidence are correct, 90% of decisions made with 90% confidence are correct, and so on and so forth. We generated calibration curves to compare the confidence-accuracy relationship for choosers and non-choosers as a function of the additional-opportunities instruction. Because calibration curves require a large number of data points, we collapsed the confidence scale into a four-point scale: 0% - 40%, 50% - 60%, 70% - 80%, 90% - 100%. We also calculated three complimentary statistics: calibration (C), over/underconfidence (O/U), and the normalized resolution index (NRI). The C statistic ranges from 0 (perfect calibration) to 1 and measures the extent to which expressed confidence deviates from accuracy. The O/U statistic varies from -1 (underconfidence) to +1 (overconfidence), with negative values indicating that the probability an eyewitness was correct exceeded his or her expressed level of confidence and positive values indicating that the eyewitness’ expressed level of confidence exceeded the probability that he or she was correct. Finally, the NRI statistic ranges from 0 (no discrimination) to 1 (perfect discrimination) and indicates how well confidence discriminates accurate from inaccurate eyewitnesses (See Brewer & Wells, 2006).
Figure 3 shows our calibration curves for choosers (top panel) and non-choosers (bottom panel) as a function of the additional-opportunities instruction and we present the complimentary calibration statistics (C, O/U, and NRI) in Table 3. There are a couple of noteworthy points that can be taken from Figure 3 and Table 3. First, replicating past research (e.g., Brewer & Wells, 2006; Wixted & Wells, 2017), confidence-accuracy calibration was generally good. This is evidenced by the positive slope of the calibration curves – which follow the ideal function (the diagonal line) quite closely – and by the fact that the identification procedures produce calibration statistics near zero (lower bound 95% CIs of .00). The one exception to this rule, non-choosers who received the additional-opportunities instruction were significantly less calibrated than non-choosers who did not receive the additional-opportunities instruction. The generally poor calibration for these participant-eyewitnesses is also evidenced by the lack of slope in the calibration curve. Moreover, the NRI statistic for these participant-eyewitnesses overlapped zero, indicating that confidence did not significantly discriminate between accurate and inaccurate non-choosers who received the additional-opportunities instruction.

The second noteworthy point coming from the calibration analysis – eyewitness-participants who received the additional-opportunities instruction and made an identification were generally underconfident (the upper-bound of the 95% CI for the O/U statistic fell below zero) and more underconfident than eyewitness-participants who made an identification in the absence of the additional-opportunities instruction (the 95% CIs for the O/U statistics did not overlap). This means that, across confidence levels, eyewitnesses who made an identification in the presence of the additional-opportunities
instruction were more likely to be correct than eyewitnesses who made an identification in the absence of the additional-opportunities instruction. This is also evident from the fact that the standard errors in the calibration plot are non-overlapping at all but the lowest levels of confidence.

**Logistic Regression.** Calibration analysis is a nice tool for assessing the confidence-accuracy relationship, but it leaves something to be desired at the applied level. Indeed, it would be ideal if we could specify by how much the odds of an accurate identification increase with each unit increase in confidence. By using confidence as a predictor in a logistic regression model, we can answer precisely this question. With each unit increase in expressed level of confidence, the odds that the identified suspect was the culprit increased 1.41 times, $B = .34, SE = .08$, Wald’s $\chi^2(1) = 19.70, p < .001, e^B = 1.41$, (95% CI [1.21, 1.63]); but, the predictive utility of eyewitness confidence did not differ as a function of the additional-opportunities instruction, $B = .08, SE = .16$, Wald’s $\chi^2(1) = .27, p = .61, e^B = 1.09$ (95% CI [0.79, 1.49]). Likewise, with each unit increase in expressed level of confidence, the odds that the rejected suspect was the culprit decreased 1.21 times, $B = .19, SE = .05$, Wald’s $\chi^2(1) = 14.75, p < .001, e^B = 1.21$ (95% CI [1.10, 1.33]); however, this effect was qualified by a trending interaction term, $B = -.18, SE = .10$, Wald’s $\chi^2(1) = 3.17, p = .08, e^B = 0.83$ (95% CI [0.68, 1.02]). Specifically, when participants were not provided with the additional-opportunities instruction, the odds that the suspect was innocent given a rejection increased 1.35 times with each unit increase in confidence, $B = .30, SE = .08$, Wald’s $\chi^2(1) = 13.82, p < .001, e^B = 1.35$ (95% CI [1.15, 1.57]). But, in the presence of the additional-opportunities instruction, confidence was a more modest predictor of accuracy for non-choosers; with each unit increase in
confidence the odds that the suspect was innocent given a rejection increased by only
1.12 times, \( B = .12, SE = .06, \chi^2(1) = 3.24, p = .07, e^B = 1.12 \) (95% CI [.99, 1.27]). This triangulates with the findings of our confidence-accuracy calibration analysis – in the presence of the additional-opportunities instruction the confidence levels expressed by non-choosers neither significantly calibrated with accuracy nor did they significantly discriminate between accurate and inaccurate non-choosers.

**Discussion**

As predicted at the outset of this paper, the additional-opportunities instruction decreased innocent suspect identifications to a greater extent than culprit identifications. We found no evidence that the standard may-or-may-not admonition has any efficacy in showup procedures. This is likely due to the fact that it is inherent in a showup procedure (but not necessarily in a lineup procedure) that the suspect may-or-may-not be the culprit. In addition, we found evidence of confidence-accuracy calibration for eyewitness identifications from showup procedures (cf. Key, Cash, Neuschatz, Price, Wetmore, & Gronlund, 2015). Confidence-accuracy calibration was less impressive among eyewitnesses who rejected suspects.

**Applied Implications and Recommendations**

The most important finding of the present research was that the additional-opportunities instruction reduced innocent-suspect identifications to a greater extent than culprit identifications. Although more research is necessary before deriving firm policy recommendations, this instruction potentially has widespread application. We found a large benefit of the additional opportunities instruction with only a slight cost. Based on an earlier conference presentation of the current work, Eisen, Smith, Olaguez, Skerritt-
Perta (2017, Experiment 3) also examined the additional-opportunities instruction. The Eisen et al. experiment was a field study with low statistical power and so the interaction between culprit-presence and the additional-opportunities instruction did not reach statistical significance. Nevertheless, the additional-opportunities instruction decreased innocent-suspect identifications from 40% to 17%, but only decreased culprit identifications from 75% to 69%.

A limitation of the present experiment is the language used in the additional-opportunities instruction. We assured participants that if they did not believe the suspect was the culprit, they would have additional opportunities to make an identification later. We used this language because we permitted eyewitnesses to view additional showups if they did not make an identification from the first showup. But, it is clear that in the real world, police officers would need to use more tentative language and instruct eyewitnesses that they might have additional opportunities to make an identification later. Eisen et al. (2017) used more tentative language in their additional-opportunities instruction (i.e., might have additional opportunities) and although the interaction between culprit presence and the additional-opportunities instruction was non-significant, the raw effect size was larger than what we found in the present manuscript. In any case, before making any firm policy recommendations, future research will need to address this issue and ensure that a more tentative instruction indicating that eyewitnesses might have additional opportunities has the same benefits as the instruction used in the present paper.

**Theoretical Implications**

The primary theoretical contribution of the current work is the demonstration that cautionary instructions can produce a better tradeoff between culprit and innocent-suspect
identifications. Further, we believe our present/absent criteria discrepancy conceptualization offers a fresh look at performance in eyewitness identification procedures. We are certainly not the first to suggest that eyewitnesses who encounter innocent suspects tend to behave differently than do eyewitness who encounter guilty suspects. Indeed, more than 30 years ago, Wells (1984, 1993) distinguished between absolute and relative judgments in an attempt to explain discrepancies between eyewitnesses in culprit-present and culprit-absent procedures. More recently, Charman et al. (2010) used the Selective Cue Integration Framework to explain why eyewitnesses who identify innocent suspects are more susceptible to post-identification feedback. We are also by no means the first to use Signal Detection Theory to explain eyewitness decision-making (e.g., Clark, 2003, 2005). But as far as we can tell, this is the first attempt to reconcile the dual-process conceptualization with signal-detection based models in the eyewitness literature. The result is a theory that predicts discrepant criteria between culprit-present and absent eyewitnesses due to strength differences in the match-to-memory experience (See Wixted & Mickes, 2010 for a similar endeavor in the basic memory literature).

The Congruent-Criteria Model. As we have noted throughout this manuscript, a classic assumption of Signal Detection Theory is that the same criteria are applied on both culprit-absent and culprit-present procedures (Green & Swets, 1966; Macmillan & Creelman, 2005). We call this the congruent-criteria model. Within the framework of this model, the only way the additional-opportunities instruction could have decreased innocent-suspect identifications to a greater extent than culprit identifications is by increasing discriminability (i.e., decreasing the overlap between the culprit and innocent
distributions in Figure 1). Could the additional-opportunities instruction have increased discriminability? Perhaps, but we can think of no plausible psychological process by which the additional-opportunities instruction would have done so. Moreover, SDT is also quite clear in its prediction that conservative instructions (e.g., the additional-opportunities instruction) should impact only criterion setting and not discriminability (Green & Swets, 1966; Macmillan & Creelman, 2005). Thus, there seems little reason to infer from a higher ROC curve that the additional-opportunities instruction increased discriminability or memory strength.

**The Discrepant-Criteria Model.** Our alternative model posits that the criteria used by participants who encounter an innocent suspect are lower than the criteria used by participants who encounter a culprit for reasons we have already discussed. Figure 1 provides the discrepant-criteria explanation for the additional-opportunities effect found in the present paper. Discriminability is equivalent in Panels A and B of Figure 1. But, Panel B displays the tendency of eyewitnesses who encounter innocent suspects to use more lenient criteria than eyewitnesses who encounter guilty suspects. The result is an increase in false alarms without a concomitant increase in hits and worse overall performance. In Panel A there is only one decision criterion that is common to both eyewitnesses who encounter innocent suspects and eyewitnesses who encounter guilty suspects. Indeed, it is our contention that the additional-opportunities instruction eliminates (or greatly decreases) the tendency for eyewitnesses who encounter innocent suspects to respond to the relatively weak match-to-memory by lowering their criteria for making an identification. The result is that the additional-opportunities instruction leads to a better tradeoff between culprit and innocent suspect identifications.
Limitations and Future Directions

The additional-opportunities instruction data we obtained are consistent with the discrepant-criteria hypothesis. But there are other theoretical explanations that could also explain this pattern of results. While we do not believe there is any reason to think that the additional-opportunities instruction increased discriminability, we cannot rule out that possibility. It is also possible that a criterial variance model could explain the additional-opportunities instruction effect. Criterial variance refers to the fact that there is variance in where eyewitnesses set their decision criteria (Smith, et al., 2017). As variance in criteria-setting decreases, performance will increase. A model that proposes a reduction in criterial variance could fit the data in the present paper. However, we can think of no mechanism by which the additional-opportunities instruction would reduce criterial variance. Maybe the additional-opportunities instruction does reduce criterial variance. But at this point we are at a loss to come up with a plausible psychological process that would lead us to expect this to happen. Accordingly, we favor the hypothesis that the additional-opportunities instruction improves performance via helping to prevent witnesses in culprit-absent conditions from setting a low criterion.

Although cautionary instructions have been examined with lineups (See Steblay, 2013), this is one of few studies to examine cautionary instructions with showup procedures (e.g., Eisen et al., 2017; Smith et al., 2014). This is an important step for the field’s understanding of how instructions work. Unlike lineups, showups do not include fillers and thus, make sorting between discriminability and response bias much more straightforward. Showups offer an important playing field for assessing the assumptions of Signal Detection Theory in the eyewitness context and for developing a better
theoretical understanding of how many variables impact eyewitness decision-making. Furthermore, additional research is needed to inform stakeholders on the best-practice recommendations for conducting showup procedures. The present work contributes to a small body of research on that topic.

Because we only used showups, we cannot be certain whether the additional-opportunities instruction effect is unique to showups or whether the same effect would also be observed with lineups. From a theoretical perspective, we see no reason to assume that the additional-opportunities instruction effect would not also apply to lineups. As with showups, a culprit-absent lineup produces a weaker match-to-memory relative to a culprit-present lineup. Hence, we would expect witnesses to use lower decision criteria when they encounter culprit-absent lineups than then they encounter culprit-present lineups. In effect, the use of the additional-opportunities instruction with a simultaneous lineup could be thought of as a backloaded simultaneous lineup.

**Final Remarks**

The additional-opportunities instruction appears to be a means by which police officers might improve performance in showup identification procedures and potentially in lineup procedures as well. Law enforcement personnel could use this instruction to reduce innocent suspect identifications with little cost to culprit identifications. Indeed, before conducting a showup procedure, officers at the Norwood Police Department in Massachusetts routinely inform eyewitnesses that they are going to be asked to view some people (William G. Brooks, Chief of Police, Norwood Police Department, Norwood, Massachusetts, USA, personal communication, March 20, 2015). The
sentiment of this instruction is similar to the additional-opportunities instruction and results in what is essentially a backloaded showup procedure.

The present/absent criteria discrepancy conceptualization is our attempt to reconcile dual-process and signal-detection approaches to eyewitness identification procedures. When researchers apply traditional Signal Detection Theory to the analysis of between-participant eyewitness data, the implicit assumption is that culprit-absent eyewitnesses construct the same fixed decision criteria as culprit-present eyewitnesses and are completely uninfluenced by the identification task (Green & Swets, 1966). This assumption runs counter to a large body of research demonstrating that people have a tendency to lower their decision criteria when match-to-memory is weak (Lampinen et al., 2014; Mansour et al., 2012; Wixted & Gaitan, 2002). Indeed, eyewitnesses who encounter culprit-absent procedures will tend to have a weaker match-to-memory experience than will eyewitnesses who encounter culprit-present procedures. Presumably, the weak match-to-memory experience afforded by a culprit-absent procedure will have the same impact as any other weak match-to-memory experience – it will lead eyewitnesses to lower their decision criteria. In this sense, eyewitnesses who encounter culprit-absent procedures tend to behave differently than eyewitnesses who encounter culprit-present procedures. But, this discrepancy between eyewitnesses who encounter culprit-absent procedures and eyewitnesses who encounter culprit-present procedures is mediated by an underlying strength variable, match-to-memory. Accordingly, we have conceptualized present/absent criteria discrepancy as a signal-detection based model.

In order for an instruction to reduce the discrepancy between the criteria set in the culprit-absent versus culprit-present conditions it has to be fashioned in a way that
capitalizes on the likely reason for the discrepancy. We tend to agree with Wixted and Gaitan (2002) and Morrell et al. (2002) that people lower their criteria when a test item produces a low match-to-memory because they do not want to miss an opportunity to get a hit on a target. This dovetails with Smith et al.’s (2014) finding with iterative showups indicating that witnesses will raise their criterion and make fewer mistaken identifications when they are surprised by the appearance of a second showup opportunity and this higher criterion perseveres throughout the remaining showups they encounter.

The theoretical explanation that we have proffered here might open up new and creative ways to improve eyewitness decision-making. There are few interventions we can think of that would increase underlying discriminability from identification procedures. Indeed, there is nothing law enforcement personnel can do to improve the conditions under which an individual witnesses a crime. But, there are many ways in which law enforcement personnel might influence response bias. To the extent that there is present/absent criteria discrepancy, we believe that it is possible to create new manipulations that will differentially impact culprit and innocent suspect identifications (as we have done here).

We believe that the present/absent criteria-discrepancy conceptualization has considerable theoretical and applied utility. In particular, the criteria-discrepancy conceptualization and the additional-opportunities instruction effect challenges the conjecture that system variables (such as instructions) can only produce a benefit of reducing mistaken identifications by producing an offsetting cost of reducing accurate identifications (Clark, 2005, 2012). This raises new possibilities for fashioning methods that might improve eyewitness identification performance.
References


### Table 1

*Percent of Suspect Identifications and Rejections as a Function of Culprit Presence and Instructions*

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Standard Instruction</th>
<th>Additional Opportunities Instruction</th>
<th>Standard and Additional Opportunities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Culprit IDs</strong></td>
<td>58% (77)</td>
<td>56% (79)</td>
<td>54% (74)</td>
<td>49% (76)</td>
</tr>
<tr>
<td><strong>Innocent IDs</strong></td>
<td>34% (77)</td>
<td>33% (76)</td>
<td>15% (75)</td>
<td>14% (76)</td>
</tr>
<tr>
<td><strong>AUC (0 – 1.00)</strong></td>
<td>.70</td>
<td></td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td><strong>pAUC (0 - .34)</strong></td>
<td>.12</td>
<td></td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td><strong>pAUC (0 - .15)</strong></td>
<td>.03</td>
<td></td>
<td>.06</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Control = control condition; Standard Instruction = may-or-may-not be the culprit; Additional Opportunities = the additional opportunities instruction; AUC = area under the curve; pAUC = partial area under the curve. Values in parentheses are the number of participants in that condition. Because ROC analysis requires a large number of observations, we collapsed the additional-opportunities and control conditions over the may-or-may-not instruction for the purpose of our ROC analysis.
Table 2

*Expressed Level of Confidence for Choosers and Non-Choosers as a Function of Accuracy and the Additional-Opportunities Instruction*

<table>
<thead>
<tr>
<th>Additional Opportunities</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
</tr>
<tr>
<td>Choosers</td>
<td></td>
</tr>
<tr>
<td>Accurate</td>
<td>65.8%</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>50.0%</td>
</tr>
<tr>
<td>$r_{pb}$</td>
<td>.33 ($p &lt; .001$)</td>
</tr>
<tr>
<td>Non-Choosers</td>
<td></td>
</tr>
<tr>
<td>Accurate</td>
<td>66.9%</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>60.7%</td>
</tr>
<tr>
<td>$r_{pb}$</td>
<td>.13 ($p = .07$)</td>
</tr>
</tbody>
</table>
Table 3

*Confidence-Accuracy Calibration for Choosers and Non-Choosers as a Function of the Additional-Opportunities Instruction*

<table>
<thead>
<tr>
<th></th>
<th>Choosers</th>
<th></th>
<th>Non-choosers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td><strong>95% CIs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>.01</td>
<td>.01</td>
<td>.00</td>
<td>.02</td>
</tr>
<tr>
<td>AOI</td>
<td>.02</td>
<td>.01</td>
<td>.00</td>
<td>.04</td>
</tr>
<tr>
<td><strong>O/U</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>-.01</td>
<td>.04</td>
<td>-.07</td>
<td>.04</td>
</tr>
<tr>
<td>AOI</td>
<td>-.15</td>
<td>.04</td>
<td>-.21</td>
<td>-.10</td>
</tr>
<tr>
<td><strong>NRI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>.08</td>
<td>.04</td>
<td>.01</td>
<td>.14</td>
</tr>
<tr>
<td>AOI</td>
<td>.16</td>
<td>.07</td>
<td>.05</td>
<td>.26</td>
</tr>
</tbody>
</table>

*Note.* C = calibration statistic; O/U = over/underconfidence; NRI = normalized resolution index; SE = standard error; 95% CIs = 95% confidence intervals; Control = control condition; AOI = additional-opportunities instruction.
Figure 1. The figure contrasts signal detection representations of the classic congruent-criteria model with our proposed discrepant-criteria model. In the discrepant-criteria model, eyewitnesses who encounter innocent suspects tend to use more lenient criteria than do eyewitnesses who encounter culprits. The result is worse performance in Panel B than in Panel A where eyewitnesses who encounter innocent suspects use the same criteria as eyewitnesses who encounter guilty suspects.
Figure 2. AOI = additional-opportunities instruction; Control = control condition; The black lines identify the “cut-off” points for the additional-opportunities instruction and control conditions, respectively. That is, these lines separate eyewitnesses who identified the suspect from those who did not. The increased diagnosticity of identification decisions that resulted from the additional-opportunities instruction is evident from comparing the distances between these lines. There is only a very small distance between these lines on the Y-axis, reflecting the fact that the additional-opportunities instruction produced only a small (nonsignificant) decrease in culprit identifications. But, the distance between the lines on the X-axis is much larger, reflecting the fact that the additional-opportunities instruction dramatically reduced innocent suspect identifications. In addition, because ROC curves in the eyewitness literature typically only include eyewitnesses who make identifications (and not those who make rejections), we have superimposed a chooser-only ROC curve in the above figure. It is evident in both representations that eyewitnesses who received the additional-opportunities instruction made more diagnostic identifications.
Figure 3. AOI = additional-opportunities instruction; Control = control condition; Error bars represent standard errors.