Title
Using Measurement Models to Understand Eyewitness Identification

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

Authors
Kaesler, Matthew
Dunn, John
Semmler, Carolyn

Publication Date
2017

Peer reviewed

Using Measurement Models to Understand Eyewitness Identification

Matthew Kaesler (matthew.kaesler@adelaide.edu.au)
School of Psychology, University of Adelaide
Adelaide, SA 5005, Australia

John Dunn (john.dunn@uwa.edu.au)
School of Psychology, University of Western Australia
Perth, WA, 6009, Australia

Carolyn Semmler (carolyn.semmler@adelaide.edu.au)
School of Psychology, University of Adelaide
Adelaide, SA 5005, Australia

Abstract

Much research effort has been expended improving police lineup procedures used in collecting eyewitness identification evidence. Sequential presentation of lineup members, in contrast to simultaneous presentation, has been posited to increase witness accuracy, though analyses based in Signal Detection Theory (SDT) have challenged these claims. A possible way to clarify the effect of presentation format on witness accuracy is to develop SDT-based measurement models, which characterise decision performance in terms of psychologically-relevant parameters, particularly discriminability and response bias. A model of the sequential lineup task was developed with a "first-above-criterion" decision rule, alongside a simultaneous model with a "maximum familiarity" decision rule. These models were fit to a corpus of data comparing simultaneous and sequential lineup performance. Results showed no difference in discriminability between the procedures and more conservative responding for the sequential lineup. Future work will examine criterion setting in the sequential lineup and model alternative decision rules.

Keywords: Eyewitness identification; Signal Detection; memory

Introduction

Typical six-person lineups used in police investigations consist of one suspect, whom police believe may be guilty of a crime, and five known-innocents called variously "fillers' or "foils", selected to resemble the suspect in specified ways (Clark, 2012). A witness may identify (ID) a person from the lineup or reject the lineup ("the person I saw is not here"), and may provide a confidence rating for their choice. In experimental mock-crime studies, a lineup is referred to as target-present (TP) if it includes the person observed by the witness at encoding (the "culprit") or target-absent (TA) if it is composed entirely of fillers.

Possible decision outcomes are expressed as rates or proportions over a series of trials. From a TP lineup, witnesses may correctly ID the culprit, incorrectly ID a foil or incorrectly reject the lineup. From a TA lineup, witnesses may correctly reject the lineup or incorrectly ID a foil, known as a false alarm.

The members of a lineup may be presented simultaneously, where the witness makes a single decision, or sequentially, where the witness makes a yes/no decision for each member before seeing the next. In experimental studies, the sequential procedure terminates once an ID is made, although variations of the procedure are used in applied settings (Horry, Palmer & Brewer, 2012). Whether presentation format affects witness accuracy has received significant research attention (Steblay, Dysart & Wells, 2011). Initial work by Lindsay and Wells (1985) found that the sequential procedure produced a marked reduction in false ID rate, a desirable outcome, and a slight reduction in correct ID rate compared to the simultaneous procedure. Numerous subsequent studies and two meta-analyses (Steblay et al., 2011) have supported this pattern of results and, although the effect has not always been found (Dobolyi & Dodson, 2013) and seems to have weakened with time (Moreland & Clark, 2016), evidence for sequential superiority has been persuasive enough for the procedure to be adopted in the United Kingdom and in many jurisdictions in the United States (Clark, 2012).

Signal Detection Theory Advantage

Recently, researchers have advocated the use of analyses based in Signal Detection Theory (SDT) to evaluate lineup procedures (Mickes, Flowe & Wixted, 2012). SDT is an approach used to analyse decision performance in a wide variety of areas in which a target, such as an enemy jet on radar or a tumor on an x-ray, must be discriminated from similar non-targets under conditions of uncertainty. In its most basic form, it characterises decision performance as resulting from two sources; *discriminability (d')*, related to how well a witness can distinguish targets from non-targets, and *response bias* or *criterion* (*c*), related to overall willingness to make a decision (MacMillan & Creelman, 2005). Claims of superior performance for the sequential procedure have been based on findings of a higher ratio of correct ID rates to false ID rates – the so called "diagnosticity ratio" – compared to the simultaneous procedure (Steblay et al., 2011). However, Mickes et al.

(2012) have shown that diagnosticity confounds discriminability and response bias. In fact, analysis of lineup data using SDT reveals that there are a range of different diagnosticity ratios associated with the same lineup procedure and that, necessarily, the ratio increases as response bias becomes more conservative, i.e. as people reject more frequently (Wixted & Mickes, 2012). SDT avoids this problem by computing all empirical or hypothetical correct ID/false ID pairs, which can then be plotted and analysed. Further, formal modelling of the task allows estimation of the entire response curve from estimated parameters – allowing tests of hypotheses about the impact of system variables such as lineup member similarity on theoretically relevant parameters.

Measurement Models

A measurement model uses theoretically-derived mathematical functions to link observed behavioural data to psychological constructs. Psychologically meaningful model parameters are estimated by fitting the model to observed data (Farrell & Lewandowsky, 2010).

Palmer and Brewer (2012) sought to address the need for formal modelling of the lineup task by fitting a SDT compound detection model (SDT-CD; Duncan, 2006) to a corpus of studies that compared simultaneous and sequential lineup data. The ‘compound’ aspect of SDT-CD refers to the fact that a lineup can be decomposed into two decision tasks; target detection (is the target present?) and target identification (if the target is present, which member is the target?). Results showed that the simultaneous and sequential lineups did not significantly differ in terms of discriminability but that the sequential procedure led to more conservative responding.

However, there were critical aspects of the analyses conducted by Palmer and Brewer (2012) that may have affected their results. First, the SDT-CD model was developed to account for simultaneous presentation of stimuli – it does not directly model the sequential procedure. For this reason, we develop and apply a formal model of the sequential procedure. Second, the best-fitting parameter values reported by Palmer and Brewer (2012) may not have been optimal as they appear not to have been estimated by an optimization procedure. For this reason, we fit the two models using a computational optimization routine. Third, given distinct models of simultaneous and sequential procedures, it is important to explore the dependence of each task on the criterion that is set by the witness. Previous research (Horry et al. 2012) has highlighted the vulnerability of the sequential task to criterion setting so we compare the simultaneous and sequential models in terms of their dependence on the decision criteria.

New Models

Both the simultaneous and sequential models assume an underlying unequal variance SDT model based on prior recognition memory research (Mickes, Wixted & Wais, 2007). Figure 1 illustrates this model for a single person lineup. Each member of a lineup is associated with a particular value of memory strength or familiarity. Foil familiarity on both TP and TA trials is modelled as a random draw from a normal distribution (dashed line in Figure 1) with mean zero and standard deviation one. Target familiarity is modeled as a random draw from a normal distribution (solid line in Figure 1), with a mean d' and standard deviation s.

A decision is made in relation to a decision criterion, c. This functions as a ‘choice threshold’ and reflects response bias. If familiarity is greater than c, the witness will identify the lineup member as the culprit. Otherwise, they reject the lineup. The greater the value of c, the more conservative is the decision and the less likely that an ID is made.

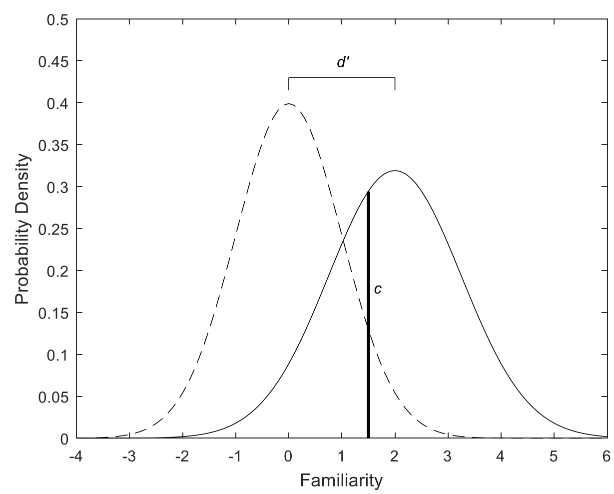
The following functions define the hit rate (h) and false alarm rate (f), where $\Phi(.)$ is the normal cumulative distribution function:

$h(c) = 1 - \Phi((c - d')/s)$

$f(c) = 1 - \Phi(c)$

For an n=1 lineup, h is the correct ID rate and f is the false ID rate.

Figure1: Basic representation of the unequal variance signal detection model



Simultaneous Model (SDT-SIM). In a simultaneous lineup, there are $n > 1$ members where typically, $n = 6$. Let $x_1, \ldots, x_n$ be the familiarity values of each member of the lineup and let m the maximum of these values. The SDT-SIM model implements the decision rule to choose member k if $m > c$ and $m = x_k$, otherwise to reject the lineup. For a TP lineup, if a choice is made and member k is the target, a correct ID has been made, otherwise, a foil ID has been made. Any selection on a TA lineup is a false alarm.

Sequential Model (SDT-SEQ). In a sequential lineup, the witness makes a decision for each member, presented in a fixed order labelled by indices, from 1 to n. Let K be a subset of these indices, such that $x_i > c$ for all $i \in K$. If K is empty then the lineup is rejected, otherwise the witness chooses member k where k is the smallest (i.e. the first) element in K. For a TP lineup, if a choice is made and member k is the target then a correct ID has been made, otherwise an incorrect ID is made. Any selection on a TA lineup is a false alarm.

Summary & Aims

Despite recent efforts to apply SDT to eyewitness identification, there has been no attempt to model the sequential lineup task taking account of the differences in the decision rule and the importance of criterion setting for response probabilities. We offer, to our knowledge, the first formal measurement model of the sequential lineup. Further, previous research did not use optimization routines to find the best fitting parameters for the data, and thus the conclusions may be different. Finally, the application of formal models in the eyewitness identification domain highlights the most important factors likely to impact on the rates of false identifications of innocent people and failure to detect the presence of a guilty suspect. In summary, the aims of the present study are:

1. To implement a formal model of the sequential lineup procedure (SDT-SEQ).
2. To reanalyse the data reported by Palmer and Brewer (2012), fitting SDT-SIM to the simultaneous lineup results and SDT-SEQ to the sequential lineup results.
3. To compare SDT-SIM and SDT-SEQ in terms of their dependence on parameter values, particularly decision criterion.

Method

Studies Analysed

A corpus of 22 studies (N = 3871) assembled by Palmer and Brewer (2012) that directly compared simultaneous and sequential presentation, making 44 data sets in all, was reanalysed. Following Steblay, et al.'s (2011) 'full diagnostic design' inclusion criteria the studies all a) manipulated both presentation format and target presence/absence, b) showed ID performance above chance levels and c) involved only adult participants.

Statistical Analyses

SDT-SIM and SDT-SEQ were fit using optimization of maximum likelihood (implemented using Matlab FMINCON function). This searches parameter space for values of d' and c that best characterise observed decision performance. We report goodness-of-fit in terms of the G^2 statistic which is a function of the maximum likelihood and distributed as chi-squared.

Statistical Considerations. Due to a lack of confidence rating data in many of the studies analysed, the standard deviation of the target distribution (s) was not estimated. Instead, the value of s was fixed to one. This is a plausible assumption in the eyewitness paradigm where each participant encodes a single study item (the culprit). The greater variance of the target distribution observed in recognition memory research may be attributed to encoding variability across a range of study items (Mickes et al., 2007).

Additionally, Palmer and Brewer (2012), following Duncan (2006), used a relative measure of criterion value with the zero point positioned between the target and lure distributions, i.e. $C = c - d'/2$. Both absolute (c) and relative (C) criteria are reported here.

Results

The new models fit the data well; SDT-SIM could not be rejected for 19 of 22 simultaneous data sets, as indicated by non-significant values of G^2. The model was rejected for data from Carlson, Gronlund and Clark (2008; Experiment 2), Lindsay and Wells (1985), and Rose, Bull and Vrij (2005). SDT-SEQ was also not rejected for 19 of 22 sequential data sets but was rejected for data from Carlson et al. (2008; Experiment 2), Lindsay and Wells (1985), and Pozzulo and Marciniak (2006). The SDT-CD model was also rejected for data from Carlson et al. (2008; Experiment 2), Lindsay and Wells (1985) and Pozzulo and Marciniak (2006), in addition to Greathouse and Kovera (2009).

Taking the parameter values estimated for each dataset, average values of d', c and C were calculated over the corpus of data, weighted according to sample size. Table 1 displays the summary results obtained by Palmer and Brewer (2012) obtained from fitting the SDT-CD model (equivalent to the SDT-SIM model) to data from both simultaneous and sequential lineups, compared to the summary results obtained by fitting the SDT-SIM model to data from both simultaneous and sequential lineups, and fitting the SDT-SEQ model to data from sequential lineups.

Table 1: Mean weighted parameter values from SDT-CD, SDT-SIM and SDT-SEQ

SDT-CD	d'	c	C
Simultaneous	1.64	-.06	-.89
Sequential	1.74	.44	-.43
SDT-SIM			
Simultaneous	1.37	1.21	.53
Sequential	1.33	1.53	.87
SDT-SEQ			
Sequential	1.40	1.55	.85

The first step in our analysis was to fit SDT-SIM to both simultaneous and sequential datasets, attempting to recover a similar pattern of results to those obtained by Palmer and Brewer (2012). Fitting SDT-SIM to all datasets produced a similar pattern of estimates to SDT-CD, with a significantly higher mean weighted C value for the sequential datasets, as indicated by a Welch two-sample weighted t-test, $t(36.42) = -3.89$, $p<.05$, and no significant difference in mean weighted *d*' values between simultaneous and sequential datasets, *t*(41.67) = .34, *p* = .73.

The second step was to compare the parameter values recovered by fitting SDT-CD and SDT-SIM/SEQ. Figures 2 and 3 plot the estimated parameter values recovered for each data set when fitting SDT-CD to both simultaneous and sequential datasets and SDT-SIM and SDT-SEQ to their respective datasets.

Figure 2. *C* vs *d'* estimates for all datasets SDT-CD

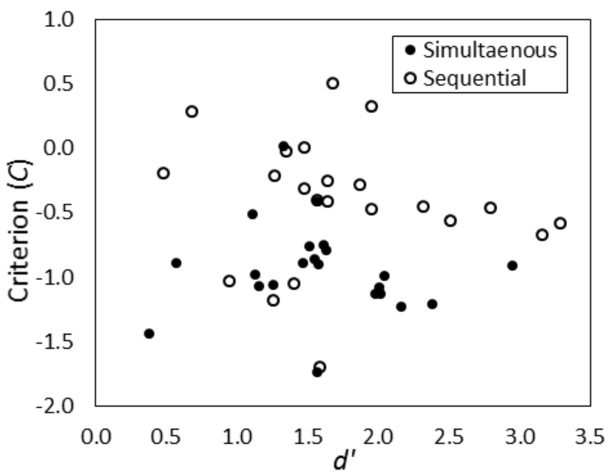
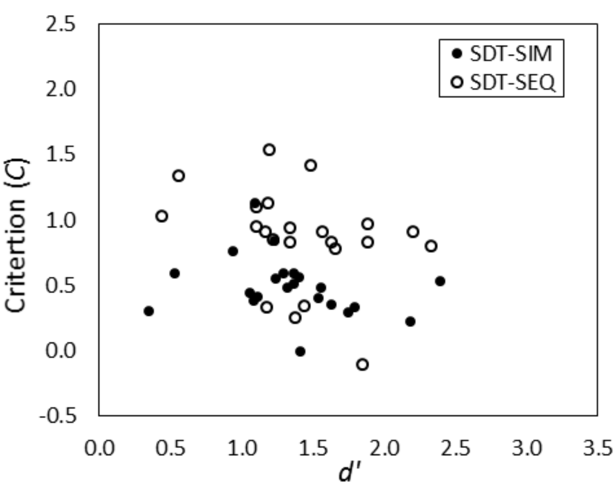


Figure 3. *C* vs *d'* estimates for all datasets SDT-SIM/SEQ



The difference in y-axis range between Figure 2 and Figure 3 indicate that fitting SDT-SIM and SDT-SEQ to their respective datasets produced higher criterion estimates compared to SDT-CD. Welch two-sample weighted *t*-tests indicated that mean weighted *C* was significantly higher for SDT-SIM, *t*(36.74) = -16.41, *p*<.05, and SDT-SEQ, *t*(38.91) = -9.87, *p*<.05, compared to mean weighted *C* from SDT-CD for simultaneous and sequential datasets respectively. There was no difference in mean weighted *d'* for SDT-SIM compared to SDT-CD as fit to simultaneous datasets, *t*(40.33) = 1.89, *p* = .06, however mean weighted *d'* for SDT-SEQ was significantly lower than SDT-CD as fit to sequential datasets *t*(33.76) = 2.11, *p*<0.5.

The final stage of the analysis was to compare *C* and *d'* values generated by fitting SDT-SIM and SDT-SEQ to their respective data types, as displayed in Figure 3. Examining Figure 3 reveals a cluster of sequential datasets with higher *C* values than the cluster of simultaneous datasets. A Welch two-sample weighted *t*-test, *t*(35.03) = -3.53, *p*<.05, indicated that the mean weighted *C* value of the sequential datasets as estimated by SDT-SEQ was significantly higher than that of the simultaneous datasets as estimated by SDT-SIM.

There are no such patterns evident relative to the horizontal axis, with *d'* values for most of the datasets clustered from approximately 1 to 2. No significant difference between the mean weighted *d*' values for the simultaneous and sequential datasets was found, *t*(41.54) = -.28, *p* = .81.

Discussion

The present study developed and fit two SDT-based formal measurement models of the simultaneous (SDT-SIM) and sequential (SDT-SEQ) eyewitness lineup task to a corpus of data collected by Palmer and Brewer (2012) using an optimization procedure, and compared the model's dependence on the parameters discriminability (*d*') and response bias (*c*) in order to better understand decision performance on the lineup task.

Fitting SDT-SIM to both simultaneous and sequential data, following Palmer and Brewer (2012), produced similar parameter estimates to those generated by fitting SDT-CD, with results suggesting that the sequential procedure encourages more conservative responding but does not differ in discriminability. Fitting SDT-SIM and SDT-SEQ to their respective data types further reinforced this pattern of results, supporting the conclusions reached by Palmer and Brewer (2012).

Compared to SDT-CD, SDT-SIM and SDT-SEQ produced higher criterion estimates for their respective data types while, while SDT-SEQ also produced lower discriminability estimates. The difference in parameter values between Palmer and Brewer (2012) and the results here is likely due to the task and fitting the models using an efficient optimization procedure rather than grid search.

While no discriminability differences were reported here or in Palmer and Brewer (2012), previous studies using ROC analysis to calculate observed discriminability from rating data have shown a discriminability advantage for the *simultaneous* lineup (Dobolyi & Dodson, 2013; Mickes et al., 2012). Results here do not necessarily contradict these findings, as simulations can generate different shaped ROC curves from different ID procedures despite holding theoretical *d'* constant between them (Rotello & Chen, 2016). The relationship between theoretical *d'* as estimated by SDT-SIM and SDT-SEQ and observed "Area Under the Curve" measures of *d'*, as used in ROC analysis, will likely be investigated in future work with new data that includes confidence ratings. Regarding theoretical discriminability, these results challenge the diagnostic feature detection model (Wixted & Mickes, 2014) that proposes a discriminability advantage for the simultaneous lineup arising from witnesses' ability to identify diagnostic features between lineup members.

The more conservative decision criterion observed on the sequential lineup can be explained by differences between its "first-above-criterion" identification rule of SDT-SEQ and the "maximum familiarity" rule of SDT-SIM. This difference becomes evident as the decision criterion is made more lenient. In the limit, the most liberal decision criterion (i.e. always choose) in the sequential lineup results in selection of the first lineup member, leading to chance performance; a correct ID rate of 1/6 for TP lineups and a false alarm rate of 1/6 for TA lineups. In contrast, the "maximum familiarity" rule of SDT-SIM means that, for the most liberal decision criterion, the witness always chooses the most familiar lineup member. If $d' > 0$, the lineup member with maximum familiarity in TP lineups is more likely than other members to be the target, leading to a correct ID rate greater than 1/6, while for TA lineups the false ID rate remains at 1/6. The effect of this difference is that in order for ID performance to be comparable between simultaneous and sequential lineups, the latter requires a more conservative decision criterion.

Based on the present findings, any performance advantage attributed to the sequential procedure is likely due to a stricter decision criterion, not improved discriminability. This suggests that changes in lineup procedure do not alter underlying memory strength. Rather, the quality of memory information available to a witness is largely determined at encoding by factors such as distance, lighting and exposure time (Maclin, Maclin & Malpass, 2001). The present findings also do not support of the proposal that performance differences are the result of procedural effects on retrieval or reconstructive memory processes taking place during a lineup decision as these are likely to affect discriminability (Ebbesen & Flowe, 2002).

Wells (2014) acknowledged the mounting body of evidence showing that any perceived sequential lineup advantage is the result of a more conservative decision criterion but contends that it is more useful in applied settings no matter the source of any performance difference. As other researchers have noted (e.g. Clark, 2012), conservative responding reduces both false IDs and correct IDs. If policy makers consider conservative responding in the lineup task desirable, such an affect could be achieved by simpler means than retraining police to administer lineups sequentially, such as instructing witnesses to be very careful in their choosing or by only counting IDs made at high confidence (Wixted & Mickes, 2012).

Limitations

Decision Strategy. The present work uses an absolute decision strategy for both the simultaneous and sequential models, despite the simultaneous lineup's long association with the so-called relative judgement strategy (Wells, 1984). Wells (1984) proposed that the increased innocent suspect ID rate of the simultaneous procedure may be due to the tendency for witnesses to compare across lineup members, selecting the one that most resembles their memory of the perpetrator *relative* to the other members, rather than comparing each lineup member directly to their memory of the perpetrator as in the absolute decision strategy (cf. Wixted & Mickes, 2014). The absolute vs. relative distinction has gained some traction in the literature and has received some empirical investigation, although the superiority of one strategy over the other has not been demonstrated (Fife, Perry & Gronlund, 2014). In line with our present approach, formal modelling of the relative decision strategy could clarify the utility of the absolute/relative distinction to understanding lineup performance.

One option for implementing relative judgement is to use the difference in familiarity between the lineup member with maximum familiarity (m) and the next-most-familiar lineup member, which seems to accord with Wells' (1984) description. The rule would be; if this difference score exceeds a criterion, then choose the lineup member with maximum familiarity, otherwise reject the lineup. We are currently developing a formal model based on this rule.

Future Directions

Criterion Shift in Sequential Lineup. The present work demonstrates that the sequential lineup decisions are critically affected by the placement of the decision criterion. A further question is whether the decision criterion may change over the course of the lineup. In an attempt to forestall such changes, Horry, Palmer and Brewer (2012) investigated the efficacy of "backloading", telling the witness that they will be viewing more photos than there are lineup members. The results indicated that the more photographs the witness was told to expect, the more conservative were their decision criteria. On non-backloaded lineups, foil choices increased in the later lineup positions. Because Horry, Palmer and Brewer (2012) fit SDT-CD to the data to generate parameter estimates, future research could explore whether these conclusions remain valid after fitting the SDT-SEQ model to these data.

Conclusion

This study presents two formal measurement models of the simultaneous and sequential lineup tasks, which were fit to a large corpus of data using computational optimization. The development of a sequential model is particularly noteworthy, it being, to our knowledge, the first of its kind in the eyewitness literature. Results show no difference in discriminability between the two procedures and a more conservative decision criterion in the sequential procedure. The models offer a means to investigate the effects of system variables on eyewitness performance in terms of theoretically relevant underlying parameters, demonstrating the value of formal modelling in applied research.

Acknowledgments

This research was funded by the Australian Government through the Australian Research Council, *Discovery Projects* DP160101048.

References

Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied, 14*(2), 118-128.

Clark, S. E. (2012). Costs and Benefits of Eyewitness Identification Reform: Psychological Science and Public Policy. *Perspectives on Psychological Science, 7*(3), 238-259.

Dobolyi, D. G., & Dodson, C. S. (2013). Eyewitness confidence in simultaneous and sequential lineups: A criterion shift account for sequential mistaken identification overconfidence. *Journal of Experimental Psychology: Applied, 19*(4), 345-357.

Duncan, M. J. (2006). *A signal detection model of compound decision tasks*. (Tech. Rep. No. TR2006–256). Toronto, ON: Defence Research and Development Canada.

Ebbesen, E. B., & Flowe, H. D. (2002). Simultaneous v. sequential lineups: What do we really know?. Retrieved December 20 2016 from https://dspace.lboro.ac.uk/2134/20167

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science, 19*(5), 329-335.

Fife, D., Perry, C., & Gronlund, S. D. (2014). Revisiting absolute and relative judgments in the WITNESS model. *Psychonomic Bulletin & Review, 21*(2), 479-487.

Horry, R., Palmer, M. A., & Brewer, N. (2012). Backloading in the sequential lineup prevents within-lineup criterion shifts that undermine eyewitness identification performance. *Journal of Experimental Psychology: Applied, 18*(4), 346-360.

Lindsay, R., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*(3), 556-564.

MacLin, O. H., MacLin, M. K., & Malpass, R. S. (2001). Race, Arousal, Attention, Exposure, and Delay: An Examination of Factors Moderating Face Recognition. *Psychology, Public Policy, and Law, 7*(1), 134-152.

MacMillan, N., Creelman., C (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, New Jersey. Lawrence Erlbaum Associates.

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*(4), 361-376.

Mickes, L., Wixted, J. T., & Wais, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review, 14*(5), 858-865.

Moreland, M. B., & Clark, S. E. (2016). Eyewitness Identification: Research, Reform, and Reversal. *Journal of Applied Research in Memory and Cognition, 5*(3), 277-283.

Palmer, M. A., & Brewer, N. (2012). Sequential lineup presentation promotes less-biased criterion setting but does not improve discriminability. *Law and Human Behavior, 36*(3), 247-255.

Pozzulo, J. D., & Marciniak, S. (2006). Comparing identification procedures when the perpetrator has changed appearance. *Psychology, Crime & Law, 12*(4), 429-438.

Rose, R. A., Bull, R., & Vrij, A. (2005). Non-biased lineup instructions do matter--A problem for older witnesses. *Psychology, Crime & Law, 11*(2), 147-159.

Rotello, C. M., & Chen, T. (2016). ROC curve analyses of eyewitness identification decisions: An analysis of the recent debate. *Cognitive Research: Principles and Implications, 1*(1), 10.

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17*(1), 99-139.

Wells, G. L. (1984). The Psychology of Lineup Identifications. *Journal of Applied Social Psychology, 14*(2), 89-103.

Wells, G. L. (2014). Eyewitness identification: Probative value, criterion shifts, and policy regarding the sequential lineup. *Current Directions in Psychological Science, 23*(1), 11-16.

Wixted, J. T., & Mickes, L. (2012). The Field of Eyewitness Memory Should Abandon Probative Value and Embrace Receiver Operating Characteristic Analysis. *Perspectives on Psychological Science, 7*(3), 275-278.

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review, 121*(2), 262-276.