Psychology, Public Policy, and Law
2006, Vol. 12, No. 2, 137–169

1076-8971/06/$12.00 DOI: 10.1037/1076-8971.12.2.137

SEQUENTIAL VS. SIMULTANEOUS LINEUPS
A Review of Methods, Data, and Theory

Dawn McQuiston-Surrett
Arizona State University

Roy S. Malpass
University of Texas at El Paso

Colin G. Tredoux
University of Cape Town

A considerable amount of empirical research has been conducted on ways to improve the eyewitness identification process, with emphasis on the use of lineups. Public policy changes are currently underway with respect to lineup procedures: Sequential lineups are being recommended to police as the best practice. This may be premature because the conditions under which sequential lineups are superior to simultaneous lineups are not well understood given the current literature: Many studies are reported with insufficient detail needed to judge the adequacy of the research design, new data show that the sequential superiority effect may vary as a function of study methodology, theoretical assumptions have not been adequately tested, and important comparisons that may rule out the ostensible superiority of the sequential lineup have not been studied. This review summarizes the literature, presents new data, and identifies the need for further empirical work before appropriately grounded recommendations as to the superiority of sequential lineups can be made.

The construction and administration of eyewitness lineups and photospreads[1] has received substantial attention in the empirical literature over the past 30 years (Doob & Kirshenbaum, 1973; Malpass, 1981; Malpass & Devine, 1983; Wells, 1993; Wells et al., 1998) and was a topic of major focus in a procedural guide for law enforcement published by the U.S. Department of Justice (Technical Working Group for Eyewitness Evidence, 1999). The two common lineup procedures used in the United States are the simultaneous lineup (SIML) and the sequential lineup (SEQL). An SIML is the traditional form of identification procedure (Wogalter, Malpass, & McQuiston, 2004), in which the witness is presented with a number of individuals and is asked whether the person who committed the crime is present. The SIML has been criticized, however, because it is said to induce witnesses to make *relative judgments* when viewing the lineup members and making a lineup decision (Dysart & Lindsay, 2001; Kneller, Memon, & Stevenage, 2001; Lindsay & Bellinger, 1999; Lindsay, Lea, Nosworthy, et al., 1991;

Dawn McQuiston-Surrett, Department of Social and Behavioral Sciences, Arizona State University; Roy S. Malpass, Department of Psychology, University of Texas at El Paso; Colin G. Tredoux, Department of Psychology, University of Cape Town, Cape Town, South Africa.

This research was supported by a grant from the National Science Foundation (SBR-9730937) to Roy S. Malpass. Portions of these data were presented in March 2006 at the annual meeting of the American Psychology-Law Society, St. Petersburg, Florida.

We thank Christian Meissner for his valuable comments on drafts of this article. Thanks also go to Otto MacLin for his contribution to this research.

Correspondence concerning this article should be addressed to Dawn McQuiston-Surrett, Department of Social and Behavioral Sciences, Arizona State University, 4701 West Thunderbird Road, Glendale, AZ 85306. E-mail: dawn.mcquiston@asu.edu

[1] We use the term “lineup” to refer to both corporeal lineups and photospreads.

Lindsay & Wells, 1985; Wells, 1984). That is, when presented with this type of lineup, witnesses are likely to compare their memory image of the perpetrator with each lineup member *relative* to the others and to choose the lineup member that most resembles the perpetrator. This strategy can be unfair toward the suspect if he or she is in fact innocent but resembles the perpetrator more than any other lineup member, particularly if the witness's recollection is imperfect.

Presenting a lineup sequentially is said to eliminate witnesses' reliance on relative judgments and to allow them to make *absolute* comparisons between their memory image of the perpetrator and each lineup member. With this procedure, lineup photographs are shown in succession and the witness is asked whether or not each photograph is that of the perpetrator. There are a number of additional procedural features of SEQLs that should be followed that differentiate them from SIMLs (Lindsay, Lea, & Fulford, 1991; Lindsay & Wells, 1985; Lindsay, 2002; Phillips, McAuliff, Kovera, & Cutler, 1999): Eyewitnesses should not know the number of photographs they will view, and often they will be led to believe that they will see two or three times the six normally shown in an SIML; they should not view any photo more than once; they should not be permitted to change a "yes" decision once it has been made; and the procedure should be conducted in a double-blind manner. Many studies have demonstrated that in comparison with SIMLs, SEQLs reduce false identifications and increase correct rejections of lineups that do not contain the perpetrator without substantially reducing the rate of correct identifications from lineups that do contain the perpetrator (Cutler & Penrod, 1988; Kneller et al., 2001; Lindsay & Bellinger, 1999; Lindsay, Lea, Nosworthy, et al., 1991; Lindsay & Wells, 1985; Melara, DeWitt-Rickards, & O'Brien, 1989; Sporer, 1993). Indeed, a recent meta-analysis (MA) examining the contrast between SIMLs and SEQLs showed evidence of a sequential superiority effect, in that SEQLs are superior under certain conditions (Steblay, Dysart, Fulero, & Lindsay, 2001). A survey indicates that the superiority of SEQLs is widely advocated in the research community: Of experts surveyed, 81% endorsed the superiority of the sequential procedure (Kassin, Tubb, Hosch, & Memon, 2001). On the basis of this body of research, recommendations are currently being made to law enforcement personnel and policymakers in a number of jurisdictions that SEQLs should be implemented to the exclusion of SIMLs (for examples, see Levi & Lindsay, 2001; Lindsay, 1999, 2002; Lindsay & Turtle, 2001; Wells, 2002), and some states, including Minnesota and New Jersey, have adopted this procedure. Media articles have highlighted this literature in pushing for a switch to SEQLs in police agencies (e.g., Kolata & Peterson, 2001).

Careful examination of the Steblay et al. (2001) MA, however, raises some important questions about the literature it summarizes and the conclusions it draws. Most significantly, the MA reported that SEQLs are superior to SIMLs for overall correct decisions (56% vs. 48%, $p < .0001$; $r = .09$). Notably, when the perpetrator-present and -absent data are separated, SEQLs are superior only when the perpetrator is absent: SEQLs are superior for minimizing false identifications of designated innocent suspects (9% vs. 27%, $p < .0001$, $r = .23$) and increasing correct lineup rejections (72% vs. 49%, $p < .0001$, $r = .25$). When the perpetrator is present, SIMLs are superior to SEQLs for securing correct identifications[2] (50%

[2]Steblay et al. (2001) note that the more realistic the experimental conditions, the smaller the

vs. 35%, respectively, $p < .0001$, $r = -.14$) and reducing false lineup rejections (26% vs. 46%, respectively, $p < .0001$, $r = -.21$). Steblay et al. (2001) collapsed over this difference and claimed that the "overall pattern of lineup accuracy supports the sequential superiority hypothesis" (p. 468), but it is more accurate to conclude that SEQLs are better than SIMLs in some respects and worse in others.

Second, inspection of the corpus of studies included in the MA shows that nearly half of the 30 tests were unpublished articles and/or conference presentations, and, importantly, the MA indicated that publication status significantly moderated the sequential superiority effect in that published articles have tended to support the effect and unpublished papers have not. This effect could be due both to submission and to reviewer biases: Manuscripts with null effects in general are less likely to be submitted and accepted for publication, and studies that do not support the sequential superiority effect may be less likely to be submitted or accepted for publication. Unpublished studies in general may also be weaker methodologically, resulting in null effects. However, it is important to note that because many of the unpublished articles included in Steblay et al. (2001) are undergraduate projects, their inclusion in the MA boosted the size of the data corpus and could possibly have led to an underestimate of the sampling and artifactual variance of effect sizes and/or to inaccurate estimates of the effects themselves.

Some proponents of MA argue against selective elimination of methodologically weak studies (e.g., Glass, 1976), touting a properly coded MA as a better test of conjectures about methodological weaknesses. We disagree with this view, particularly because recent policy initiatives around eyewitness identification (Technical Working Group for Eyewitness Evidence, 1999) will increasingly put eyewitness research itself in the dock. There is more to the evaluation of the details of methods, procedure, and their reporting than the correlations between them and other variables in the literature, as the audience for the studies and their policy implications go far beyond that of other researchers. For example, it could be difficult to convince legislators, lawyers, judges, and juries when fully 40% of the knowledge about SEQLs derives from unpublished undergraduate projects and academic colloquia and has not been subject to peer review.

Because a goal of research in this domain is to change public policy (Lindsay, 1999), one possibility is to follow a different standard, for instance, a "Daubert-like" approach (*Daubert v. Merrell Dow Pharmaceuticals*, 1993) and thus to exclude papers or other evidence that has not been subjected to peer review. For example, in a recently published MA examining the effects of stress on eyewitness memory, Deffenbacher, Bornstein, Penrod, and McGorty (2004) wrote, "No unpublished studies were included, because the legal standards for proffered scientific testimony established by the U.S. Supreme Court in *Daubert* . . . have strengthened the preference by the legal system for meta-analytic conclusions based on a body of well conceived, well executed, and easily retrievable studies" (p. 692). Thus, the issue of the selection of studies for an MA ultimately goes beyond the comparison of SEQLs and SIMLs. However, we do not say that the Steblay et al. (2001) MA is at fault for including so many pieces of unpublished undergraduate research. Indeed, Steblay et al. are correct in terms of the standard

---

difference in correct identifications for simultaneous vs. sequential lineup procedures (more realistic conditions produce $r = -.11$, less realistic conditions produce $r = -.30$).

meta-analytic approach by testing for, and declaring, a strong moderator effect in this respect. The question is whether the standard approach might involve dangers of the kind noted by Deffenbacher et al. (2004) when translated into the realm of policy making rather than research synthesis.

Third, publication of the Steblay et al. (2001) MA highlighted—in our opinion—the relatively small size of this research literature (30 comparisons are reported). Only a limited range of possible variables with respect to the SIML/SEQL comparison have actually been studied, and some findings have not been replicated. This is not surprising; small literatures are more likely to show uneven coverage of the possible topics and questions that can be studied in a research field. However, the small size of the literature seems to have led to a fundamental confound: Whereas SIML versus SEQL studies are generally framed as the comparison of a lineup that is administered by presenting lineup members to the witness, either one at a time or all at once, these studies have introduced more changes to the classic SIML than simply the presentation mode in order to produce an SEQL (i.e., declaration of the number of lineup members to be shown; number of identification questions the witness is asked). These changes would have been better treated as experimental variables at some stage in the development of the literature but remain unexplored and are potential sources of confound in the study of SIMLs versus SEQLs. A discussion of what is missing from the study space concerning SIMLs versus SEQLs is presented later in this article.

It is also important to note that a great many of the studies included in the SIML/SEQL MA were conducted in a single laboratory, which may raise questions about the independence of studies selected for review. This also suggests that the estimates of the SEQL advantage arrived at in the MA could be upwardly biased. As a check on this assertion, we recalculated some meta-analytic statistics for the corpus of studies analyzed in the MA, using as a moderator variable the identity of the author who runs the laboratory we refer to above (R. C. L. Lindsay). That is, we partitioned the set of individual studies referenced in the Steblay et al. (2001) paper into two sets: one in which Lindsay's name appears as either the first or subsidiary researcher ($n = 11$) and one in which his name does not appear on the author list ($n = 13$).[3] Steblay et al. reported a number of study statistics, but for the purpose of succinctness we repeat their calculations for three key statistics, namely, correct identifications in perpetrator-present (PP) lineups (hits), false positive identifications in perpetrator-absent (PA) lineups (false alarms), and overall identification accuracy. In the case of hits, we find .55 versus .60 (SEQL vs. SIML) for the single laboratory set and .28 versus .46 for the second set. The second difference is statistically significant ($r = .17$, $z = 4.77$, $p < .01$), but the first is not. In the case of false alarms, we find .25 versus .58 (SEQL vs. SIML) for the single laboratory set and .29 versus .46 for the second set. Both differences are statistically significant ($r = .31$, $z = 9.58$, $p < .01$, for the first, and $r = .21$, $z = 4.8$, $p < .01$, for the second). Finally, the difference in the overall number of accurate decisions in the case of the single laboratory set is .67 versus .46 (SEQL vs. SIML) for the single laboratory set and .51 versus .50 for the second set. The first of these is statistically significant ($r = .26$, $z = 7.48$, $p < .01$), but the second is not.

[3]Only studies that provided data on correct identifications, false alarms, and overall identification accuracy are included in this analysis.

This simple analysis shows that studies from one laboratory produce results of no statistically detectable difference between SEQLs and SIMLs in the rate of correct positive identifications, whereas studies from other laboratories suggest that SIMLs are 1.64 times more likely to lead to correct positive identifications than are SEQLs. More significantly, those studies not conducted in the single laboratory almost perfectly balance reductions in the false positive rate (from .46 to .29, a ratio of 1.59, achieved with SEQLs), with reductions in the positive identification rate (from .46 to .28, also achieved with SEQLs). Bluntly put, outside of the corpus of published studies emanating from the single laboratory, there is no evidence that SEQLs are in overall terms superior to SIMLs. We emphasize strongly that this is not an *ad hominem* argument and believe that several interpretations of this finding can be made. First, there may be procedural differences between Lindsay's studies and others. Second, it is possible that Lindsay's laboratory is more careful and precise in administering the essential procedural elements of their studies than are others. These interpretations are speculative, however, thus, a fuller (moderator) analysis of this laboratory effect can be found later in this article.

In general, MAs provide a computational summary of a research area and do not necessarily critique a body of literature or examine how comprehensive a literature is in covering important design variables or the quality/methodological strength of the corpus of studies (Beaman, 1991; Cooper & Hedges, 1994; Cooper & Rosenthal, 1980). The risk is that effect sizes derived from an MA can be misleading if they are based on the results of individual studies that are methodologically weak or whose findings are ambiguous or questionable. Indeed, the selection of studies for inclusion in Steblay et al. (2001) resurrects questions about the places of methodological evaluation and peer review in MAs. On the basis of our review, we believe that the literature concerning SIMLs versus SEQLs may be underdeveloped in some important ways and that important sources of confound have yet to be addressed. The MA may be a good summary of the findings in the existing literature, but because of the paucity of research making up the total corpus and other potential limitations, the conclusions that can be drawn about the differences between SEQL and SIML presentations per se may be limited. Some may argue that the fact that the literature is made up of studies of quite different kinds, perhaps with the methodological limitations we allege, strengthens rather than weakens the conclusions of the MA. That is, the fact that there are differences despite the "random noise" means that there is convergent validity among studies with different types of errors, and this can be viewed as lending support to the MA conclusions. We do not believe that this is the case: Steblay et al. identified multiple sources of systematic variance in their MA, and some of these substantially reduced the SIML–SEQL difference effect size. In the case of many of the other methodological limitations we identify, both in the discussion above and that which follows, there is not enough detail in the articles to identify whether the variations in effect size are systematic.

Because of the potential implications of the literature for developments in public policy and because the MA has been put forth to represent the findings of this body of literature, we believe that a deeper analysis of all published and unpublished research in this area is in order. We assert that a thorough review must include careful consideration of every study that Steblay et al. (2001) have defined as the corpus of studies making up this literature, whether published or unpublished. In what follows, we attempt to provide a methodological and statistical summary and evaluation of the extant literature on SIMLs versus SEQLs.

The Current Review and Analysis

A literature search was conducted in an effort to obtain a complete and up-to-date reference list on the SIML–SEQL comparison, including the reference list from the Steblay et al. (2001) MA. Our selection criteria for including articles in this review (beyond that in the MA) were the following: (a) works must come from refereed sources, including published papers, conference papers, and doctoral dissertations, and (b) studies must provide a statistical test of the SIML versus SEQL comparison. We searched several electronic databases (e.g., PsycINFO, Dissertation Abstracts International, Social Sciences Citation Index) and relevant conference programs spanning the past 10 years (e.g., American Psychological Association and numerous divisions therein; American Psychological Society) and contacted relevant researchers to gain additional information about published and unpublished sources. We identified 37 articles in the literature containing 45 experiments comparing SIMLs with SEQLs. Of these 37, a total of 17 were published and 20 were unpublished (consisting of 11 undergraduate theses, 1 master's thesis, 2 doctoral dissertations, 2 conference presentations, and 4 unpublished manuscripts). We obtained full-text copies of all articles except for 10 experiment write-ups, in which summaries of the studies were made available to us. Specific methodological attributes of each experiment were then coded (see Table 1). Table 2 contains a list and description of the attributes.

We separated the data derived from our literature review into three main sections for purposes of analysis. We first present a descriptive analysis concerning research design, methodological, and data interpretation issues. Next, we present the results of a moderator analysis that extends the search for moderators of the sequential superiority effect. We then turn to an analysis of the most commonly advanced theoretical explanation of the ostensible advantage that SEQLs have over SIMLs: relative versus absolute judgments.

Descriptive Analysis

Considering Important Procedures, Design Variables, Methodology, and Underreporting

A range of variables can weaken the reliability and validity of experimental findings in psychological research (e.g., Cook & Campbell, 1979). Studies in eyewitness identification are no exception (Wells & Luus, 1990); in our review of the published and unpublished literature we discovered some important issues that often fail to be considered in the design of SIML versus SEQL studies. We also found a surprising number of important details relevant to methods and procedures used that were not reported in the published articles,[4] making it difficult to evaluate aspects of many studies and to interpret results clearly. The following details the results of our analysis regarding these methodology issues.

[4]We included only published articles in our analysis of underreported details because we would not expect unpublished papers (e.g., conference presentation handouts) to include the same amount of detail as an article published in a scientific journal. In addition, a large portion of the unpublished papers that make up the corpus of this literature as reported in Steblay et al. (2001) are undergraduate theses, and only summaries of these papers are available.

Perpetrator–suspect similarity. A PA lineup is often constructed in eyewitness research by substituting a similar-looking person for the perpetrator—a procedure that is easy to use in the laboratory but not necessarily consistent with the research recommending description-matched lineups (Wells, Seelau, Rydell, & Luus, 1994). This strategy also allows researchers to study the "worst-case scenario" in the laboratory, modeling the situation in which the police suspect matches the offender's physical description yet is in fact innocent (Lindsay & Wells, 1985). A potential problem with this strategy is that suspects highly similar to the perpetrator will draw more false identifications and more identifications than other lineup members, depending on the filler selection method used. And, if some fillers are sufficiently dissimilar that they are easily eliminated as viable lineup choices, PA suspect selection is facilitated. Thirty-six percent of published studies and 88% of unpublished studies reported using high-similarity suspect replacements, whereas 45% of published studies and 6% of unpublished studies reported using suspects "moderately similar" to the perpetrator. Descriptions of the procedures used to determine high versus moderate similarity and quantitative evidence are often not provided, so the empirical basis for the claim is seldom known beyond the investigator's impressions. Likewise, the faces in the lineup are seldom published, so others cannot do the similarity scaling. An important matter for interpretation is to determine the difference between "highly similar" and "moderately similar." Of the published literature, 26% did not report information about the suspect's physical appearance.

Lineup construction method. Research demonstrates that the method of lineup filler selection used can impact identification performance (e.g., Tunnicliff & Clark, 2000). For example, suspect-matched lineups can result in bias or suggestiveness toward the suspect (Laughery, Jensen, & Wogalter, 1988; Marwitz & Wogalter, 1988; Wogalter, Marwitz, & Leonard, 1992), whereas description-matched lineups are said to allow for all lineup members to vary on general physical characteristics so as to not make the lineup task impossible for an eyewitness who has a good memory of the perpetrator (Wells et al., 1994). Yet we do not know the relative frequencies in real criminal identification cases of suspects being nominated on the basis of a verbal description or on the basis of information unrelated to physical appearance. In the latter case, filler selection based on a variation of suspect matching (Wogalter et al., 1992) might be more appropriate. Filler selection was based on a match to physical similarity in 58% of the published studies and in 26% of the unpublished studies and was based on a match to physical descriptions in 13% of the published and 3% of the unpublished studies. Experimenters selected the fillers (i.e., no pilot testing) in 44% of the published and 10% of the unpublished studies.

Double-blind lineup administration. Double-blind testing is standard procedure in medical and pharmaceutical research and is used in some areas of experimental psychology to prevent inadvertent communication of information to research participants about critical aspects of the research. Blind administration is an important step to take in any eyewitness identification procedure, and there is recent evidence that SEQLs are inherently more susceptible to subtle and inadvertent distinguishing of the suspect (Haw & Fisher, 2004; Parker, Tredoux, & Nunez, 2000; Phillips et al., 1999). Indeed, it has been argued that SEQLs should be used only under these conditions (Levi & Lindsay, 2001; Wells et al., 1998).

(*text continues on page 148*)

Table 1
Characteristics of Sequential vs. Simultaneous Lineup Studies

Study	Culprit presence or absence	Number of culprits/targets	Lineup construction method	Culprit–suspect similarity	Number of designated suspects in culprit-absent lineup
*Dormer (1983) Unpublished undergraduate thesis	Present and absent	1	Pilot study (similarity ratings)	High[b]	1
*Hannaford (1985) Unpublished undergraduate thesis	Present and absent	1	—[a]	High[b]	1
*Lindsay & Wells (1985) Published	Present and absent	1	Constructed by experimenter (matched target appearance)	High	1
*Cutler & Penrod (1988), Exp. 1 Published	Present and absent	1	Constructed by experimenter (matched target appearance)[b]	High[b]	All-suspect lineup
*Cutler & Penrod (1988), Exp. 2 Published	Present and absent	1	Constructed by experimenter (matched target appearance)[b]	High	1
*Melara, DeWitt-Rickards, & O'Brien (1989), Exp. 1 Published	Present and absent	1	Constructed by police officers (matched target appearance)	Moderate	Cannot determine from analysis
*Lindsay, Lea, & Fulford (1991), Exp. 1 Published	Present and absent	1	Constructed by experimenter (matched target appearance)	Moderate	1
*Lindsay, Lea, & Fulford (1991), Exp. 2 Published	Absent only	1	Constructed by experimenter (mismatched target appearance)	Moderate	1
*Lindsay, Lea, & Fulford (1991), Exp. 3 Published	Absent only	1	—[a]	Moderate	1
*Lindsay, Lea, Nosworthy, Fulford, Hector, LeVan, & Seabrook (1991), Exp. 1 Published	Present and absent	1	Constructed by experimenter (matched target appearance)	Moderate	1
*Lindsay, Lea, Nosworthy, Fulford, Hector, LeVan, & Seabrook (1991), Exp. 2 Published	Present and absent	1	—[a]	—[a]	1
*Lindsay, Lea, Nosworthy, Fulford, Hector, LeVan, & Seabrook (1991), Exp. 3 Published	Absent only	1	Constructed by experimenter (matched/mismatched target appearance)	High	1
*Lindsay, Lea, Nosworthy, Fulford, Hector, LeVan, & Seabrook (1991), Exp. 4 Published	Absent only	1	—[a]	—[a]	1
*Lindsay, Lea, Nosworthy, Fulford, Hector, LeVan, & Seabrook (1991), Exp. 5 Published	Absent only	1	Constructed by experimenter (mismatched target appearance)	High	1
*Parker & Ryan (1993) Published	Present and absent	2 (between-Ss)	Pilot study (similarity ratings)	Moderate	1
*Sporer (1993) Published	Present and absent	1	Pilot study (similarity ratings)	High	Cannot determine from analysis
*Jacob (1994) Unpublished undergraduate thesis	Present and absent	0–3 (within-Ss)	—[a]	High	1
*Lindsay, Martin, & Webber (1994), Exp. 3 Published	Absent only	1	Pilot study (similarity ratings)	Moderate	1
*Rombough (1994) Unpublished undergraduate thesis	Present and absent	1	Pilot study (similarity ratings)	High[b]	1
*Smyth (1994) Unpublished undergraduate thesis	Present and absent	1	Pilot study (similarity ratings)	High[b]	1
*Varrette (1994) Unpublished undergraduate thesis	Present and absent	1	—[a]	High[b]	1

Lineup manipulation check	Lineup photos counter-balanced	Blind lineup testing	SEQL method	Backloading of SEQLs	SEQL stopping rule	Treatment of multiple IDs
No[b]	No	No[b]	Photographs[b]	BL[b]	All photos presented[b]	—[a]
No[b]	Yes	No[b]	Photographs[b]	BL[b]	All photos presented[b]	—[a]
No[b]	Yes	No	Photographs	BL	All photos presented	Foil ID
No[b]	No	No, but photos not physically handled	Video and projected onto screen	Non-BL	Instructed to make one ID	No multiple IDs
No[b]	No	No, but photos not physically handled	Projected on screen	BL	Instructed to make one ID	No multiple IDs
—[a]	Yes	Yes	Photographs	Non-BL	—[a]	—[a]
Yes	No	No	—[a]	Non-BL	All photos presented	—[a]
—[a]	No	No	—[a]	BL	All photos presented	—[a]
—[a]	No	No, but photos not physically handled	Projected onto screen	BL	All photos presented	—[a]
—[a]	No	No	Photographs[b]	BL[b]	All photos presented[b]	—[a]
—[a]	No	No	Photographs[b]	BL[b]	All photos presented[b]	—[a]
—[a]	No	No	Photographs[b]	BL[b]	All photos presented[b]	—[a]
—[a]	No	No	Photographs[b]	BL[b]	All photos presented[b]	—[a]
—[a]	No	No	Photographs[b]	BL[b]	All photos presented[b]	—[a]
Yes	Yes	No	Photographs[b]	Non-BL	Instructed to make one ID	Foil ID
—[a]	Yes	No	Photographs	—[a]	Allowed multiple IDs	Analyzed first ID
No[b]	Yes	No[b]	Photographs	BL[b]	All photos presented[b]	—[a]
—[a]	No	No	Photographs	BL	All photos presented[b]	—[a]
No[b]	Yes	No[b]	Photographs[b]	Non-BL	Instructed to make one choice	Foil ID
No[b]	Yes	No[b]	Photographs[b]	Non-BL	Instructed to make one choice	Foil ID
No[b]	No	No[b]	Photographs	BL[b]	All photos presented[b]	—[a]

(*table continues*)

Table 1 (*continued*)

Study	Culprit presence or absence	Number of culprits/targets	Lineup construction method	Culprit–suspect similarity	Number of designated suspects in culprit-absent lineup
*Lindsay, Pozzulo, Craig, Lee, & Corber (1997), Exp. 1 Published	Present and absent	1	Constructed by experimenter (matched target appearance)	—[a]	All-suspect lineup
*Lindsay, Pozzulo, Craig, Lee, & Corber (1997), Exp. 2 Published	Present only	1	Constructed by experimenter (matched target appearance)	[No culprit, absent lineup]	(No culprit-absent lineup)
*Martins (1996) Unpublished undergraduate thesis	Present only	—[a]	—[a]	[No culprit, absent lineup]	—[a]
*Vanderwal (1996) Unpublished undergraduate thesis	Present and absent	0–3 (within-Ss)	—[a]	High	1[b]
*Laldin (1997) Unpublished undergraduate thesis	Present and absent	0–3 (within-Ss)	—[a]	High[b]	—[a]
*Bellinger (1995) Unpublished undergraduate thesis (published as Lindsay & Bellinger, 1999; culprit-absent data from unpublished paper included in the meta-analysis)	Absent only	4 (between-Ss)	—[a]	High	1
*Newman (1998) Unpublished undergraduate thesis	Present and absent	1	—[a]	High[b]	1
*Blank & Krahe (2000) Conference presentation	Present only	3	Constructed by experimenter (matched target description)	[No culprit, absent lineup]	(No culprit-absent lineup)
*Parker, Tredoux & Nunez (2000) Unpublished manuscript	Absent only	2 (between-Ss)	Pilot study (similarity ratings)	High	1
*Kneller, Memon, & Stevenage (2001) Published	Present and absent	1	Pilot study (similarity ratings)	High	1
*Willis (1990) Unpublished doctoral thesis	Present only	1	—[a]	(No culprit, absent lineup)	(No culprit-absent lineup)
Cox (1992) Unpublished doctoral thesis	Present and absent	1	Pilot study (cluster analysis)	High	1 (no distinction between suspect and foil IDs in analysis)
Yarmey & Morris (1998) Published	Absent only	1	Pilot study (similarity ratings)	High	1
Phillips, McAuliff, Kovera, & Cutler (1999) Published	Present and absent	2 (event had 2 culprits)	Pilot study (similarity ratings)	High	1
Dysart (1999), Exp. 1 Unpublished master's thesis	Present and absent	1	(Lineup 1) constructed by experimenter (matched target appearance; (Lineup 2) pilot study (similarity ratings)	High	All-suspect lineup
MacLin & Malpass (2001) Unpublished manuscript	Present and absent	1	Pilot study	High	1
McQuiston, Salinas, Villegas, Garcia, & Malpass (2001) Conference presentation	Present and absent	2 (between-Ss)	Pilot study (similarity ratings)	High	1
Shapiro & Hiatt (2003) Unpublished manuscript	Present and absent	1	Constructed by experimenter (matched target appearance/description)	—[a]	All-suspect lineup
Dysart & Lindsay (2001) Published	Present and absent	1	Constructed by experimenter (matched target appearance)	Moderate	1
Gaitens, Zimmerman, McQuiston, & Malpass (2002) Conference presentation	Absent only	1	Pilot study (determined whether anyone stood out)	Moderate	All-suspect lineup

Lineup manipulation check	Lineup photos counter-balanced	Blind lineup testing	SEQL method	Backloading of SEQLs	SEQL stopping rule	Treatment of multiple IDs
No	No	No, but photos not physically handled	Photographs in booklet	BL	All photos presented	Excluded from analysis
No	Yes	No, but photos not physically handled	Photographs in booklet	BL	All photos presented	Excluded from analysis
No[b]	—[a]	No[b]	Photographs[b]	BL[b]	All photos presented[b]	—[a]
No[b]	Yes	No[b]	Photographs	BL[b]	All photos presented[b]	—[a]
No[b]	—[a]	No[b]	Photographs	BL	All photos presented[b]	—[a]
No	No	No	Photographs	BL	All photos presented	—[a]
No[b]	—[a]	No[b]	Photographs	BL[b]	All photos presented[b]	—[a]
No[b]	Yes	No	Photographs	BL	Sequence stopped after ID made	No multiple IDs
Yes	Yes	No, but photos not physically handled	Computer screen	BL	All photos presented	Data analyzed both with and without multiple IDs
No[b]	Yes	No	Photographs	BL	Sequence stopped after ID made	No multiple IDs
Yes	Yes	No, but photos not physically handled	Projected onto screen	BL[b]	Sequence stopped after ID made	No multiple IDs
—[a]	No	No	Video	Non-BL	—[a]	—[a]
—[a]	Yes	No	Photographs	BL	Sequence stopped after ID made	No multiple IDs
Yes[b]	No	Yes (in some conditions)	Photographs	BL	Sequence stopped after ID made	No multiple IDs
Yes[b]	No	No	Photographs	BL	All photos presented	No multiple IDs[b]
No	Yes	No, but photos not physically handled	Computer screen	BL	Ss told lineup would end upon ID	No multiple IDs
No	Yes	No	Photographs	BL	All photos presented	Excluded from analysis
No[b]	Yes	No, but photos not physically handled	Computer screen	BL	Depended on which condition (2 SEQL procedures)	No multiple IDs[b]
Yes[b]	No	No	Photographs	BL	All photos presented[b]	No multiple IDs[b]
Yes	Yes	No, but photos not physically handled	Computer screen	Non-BL	All photos presented	Excluded from analysis

(*table continues*)

Table 1 (*continued*)

Study	Culprit presence or absence	Number of culprits/targets	Lineup construction method	Culprit–suspect similarity	Number of designated suspects in culprit-absent lineup
Memon & Bartlett (2002) Published	Present only	1	Pilot study (matched target description)	(No culprit-absent lineup)	(No culprit-absent lineup)
Melssner (2002), Exp. 1 Published	Present and absent	3 (between-Ss)	Pilot study (matched target description)	Moderate	All-suspect lineup
Memon & Gabbert (2003a) Published	Present and absent	1	Pilot study (matched target description)	—[a]	1
Memon & Gabbert (2003b) Published	Present only	1	Pilot study (matched target description)	(No culprit-absent lineup)	(No culprit-absent lineup)

Note. SEQL = sequential lineup; BL = backloading; Ss = subjects.
[a]Information not available in the original paper and not provided in response to our inquiry to the author(s). [b]Information not available in the original paper but was provided in response to our inquiry to the author(s).
*This study was included in the Steblay et al. (2001) meta-analysis.

Eight percent of the published studies and 75% of the unpublished studies reported using true double-blind procedures. Thirty-six percent of published and 25% of unpublished studies used methods in which lineup presentation was out of the hands of the experimenter (e.g., computerized administration). Because of the limited investigation on this important topic, there is a need for further research addressing claims that blind testing is an essential aspect of the SEQL procedure. Such research can clarify the conditions under which blind lineup administration is advantageous.

Stimulus sampling. In eyewitness studies, one individual is typically designated as the target person, and recognition of that person is later tested. A potential problem with this approach is that sampling only one person from the chosen category of interest (e.g., a particular ethnic group and/or gender) may not adequately represent that category, threatening construct and external validity (Cook & Campbell, 1979; Wells & Windschitl, 1999). To the extent that a particular face is processed and/or remembered differently from other faces (e.g., based on distinctiveness, attractiveness, etc.), it is important to sample adequately from a category. Eighty-eight percent of published studies and 60% of unpublished studies reported testing recognition of only one target face.

Position effects/counterbalancing. Counterbalancing, as all students of experimental design are taught, is an important laboratory safeguard and can be used to control for position effects. Order and position effects can contaminate the results of an experiment or even produce results that are misleading. In the case of experimental studies of lineups, the use of counterbalancing is similarly important. This is particularly clear with regard to the order in which lineup members are arranged, both for SEQLs and SIMLs. Some authors have reported position, or contrast, effects for different lineup orders in SIMLs (Gonzalez, Davis, & Ellsworth, 1995; Sporer, 1993); counterbalancing can be expected to remediate or alleviate these problems. The failure to counterbalance foil order, on the other hand, can result in a very suggestive lineup structure, especially if there is any “leakage” within the experimental design (e.g., some participants perhaps telling others which lineup member in the sequence they chose). A counterbalanced structure would remediate this threat, yet only 32% of published and 60%

Lineup manipulation check	Lineup photos counter-balanced	Blind lineup testing	SEQL method	Backloading of SEQLs	SEQL stopping rule	Treatment of multiple IDs
Yes	No	No, but photos not physically handled	Photographs in booklet	BL	—[a]	—[a]
Yes	Yes	No, but photos not physically handled	Projected onto screen	BL	All photos presented	No multiple IDs
Cannot determine	No	No, but photos not physically handled	Photographs in booklet	BL	Ss told lineup would end upon ID	No multiple IDs
Cannot determine	No	No, but photos not physically handled	Photographs in booklet	BL	Ss told lineup would end upon ID	No multiple IDs

of unpublished studies reported using counterbalancing methods. There are many further problems that could beset experimental studies of lineups that are not counterbalanced that we do not list here; our point is that counterbalancing is a basic aspect of laboratory practice, and it should not have been jettisoned in studies that use lineups.

Manipulation check. It is important to establish empirically and objectively the "fairness" of lineups created and used in experimental research prior to their implementation; a lineup constructed and perceived to be fair (Malpass, Tredoux, & McQuiston-Surrett, in press) by one person may not be judged so by other people. Checking the fairness of a lineup can be achieved using methods such as a mock witness test or similarity ratings (Doob & Kirshenbaum, 1973; Malpass, 1981; Malpass & Devine, 1983; Malpass et al., in press; Tredoux, 1998, 1999). Seventy-six percent of published studies did not report whether a manipulation check was used on lineups, thus, we do not know to what extent findings of a sequential superiority effect is moderated by the fairness of the lineup. This is also something that the Steblay et al. (2001) MA cannot address, as very few studies report data on lineup fairness.

Backloading instructions. An important lineup instruction given to SEQL participants pertains to their awareness of *backloading* of the sequence, that is, the practice of presenting the same fillers that are presented to witnesses viewing the SIML but "loading" (even if only by implication) some additional number of fillers behind the initial set. The purpose of backloading is to withhold from the witness the exact number of faces in the sequence, and this is a central rule for retaining the advantage of the SEQL (Lindsay, Lea, & Fulford, 1991). Twenty-four percent of the published studies did not report whether participants were aware how many photographs they would view. In 16% of the studies, participants knew how many photos they would view in the SEQL condition.

Stopping rule instructions. The "stopping rule" informs participants who view an SEQL whether or not the display of photos will conclude on the first positive identification made. In the original formulation of the SEQL, the termination of the lineup at the point of the first identification was thought to be a key element (Lindsay & Wells, 1985). Not all studies have followed the original rule,

Table 2
Study Characteristics Catalogued in Table 1

Study characteristic	Description
Perpetrator presence or absence	Whether the study tested identification using perpetrator present or absent lineups or both
Number of perpetrators	The number of perpetrators used in the study and whether the use of multiple perpetrators was a within- or between-participants variable
Lineup construction method	Method by which lineups were constructed
Perpetrator–suspect similarity	The degree of similarity between the designated suspect (if there was one) and the perpetrator
Number of designated suspects in perpetrator-absent lineup	Whether the perpetrator-absent lineup contained only one designated suspect or all members of the lineup were considered suspects
Lineup manipulation check	Whether methods were used to determine lineup fairness prior to implementation in research
Lineup photographs counterbalanced	Whether lineup photographs were randomized or counterbalanced for presentation
Blind lineup testing	Whether the lineup administrator was aware of the identity or position of the target person in the lineup
Sequential lineup administration method	The procedure used to administer sequential lineups
Backloading of sequential lineups	Whether participants were led to expect more lineup photographs than would actually be shown by the administrator in the sequential condition
Sequential lineup stopping rule	What instructions were given to participants indicating that the sequential photographic display would end upon the first positive identification made
Treatment of multiple identifications in analysis	Whether multiple identifications were made or allowed and, if so, how they were handled in the data analysis

and it is important to know exactly what rule a particular study has used. Witnesses may have a high degree of variability in their beliefs about the structure of the task as a function of the type of instruction concerning when the lineup will end. Of the published studies, 36% did not report what instruction was given to SEQL participants with respect to the stopping rule. Twelve percent of the published studies informed participants that the lineup would conclude when an identification was made, 20% instructed participants to make only one identification, and 56% presented all photographs to participants, whether or not an identification was made.

Omissions in Research Design and Control Variables

A number of variables are confounded in the research designs that form this corpus of studies. These are factors that are part of the package of variables comprising SEQLs but that have not, until recently, been added to SIML and isolated as design elements (Zimmerman, Malpass, & MacLin, 2006). Because of this omission, it is difficult to know whether it is these variables, taken alone or in combination, that account for the contrasts found between SIML and SEQL or whether it is the sequential presentation form per se. This is an important distinction, because if the sequential presentation form is the important element, this implies that the relative–absolute contrast is significant. If the contrast depends on the other factors, the judgment theory may take on less importance. Two factors of particular interest are backloading of additional lineup members and asking individual questions for each lineup member.

Backloading. Including additional faces to follow the usual six can be expected to produce conservative witnesses' responses because they expect to see many more faces and, depending on the specific instruction, may be reluctant to make an identification early when there are so many more faces yet to be seen. If so, then this factor is confounded between the two lineup forms and might be the basis for the reduced identification rates that are seen in both PP and PA lineups. This possibility can easily be tested by comparing a backloaded SEQL procedure with an SIML in which witnesses are also shown a backloaded set of faces (e.g., a total of 18, in three groups of 6).

One versus multiple questions. The procedure of asking the identification question ("Is this the person who . . .?") of each face in SEQLs could be applied to SIMLs as well. Whereas traditional SIML instructions ask the global question, "Is the person you saw . . . in the lineup?" and whereas this question is generally used for each face in the SEQL, one could also ask the same question of each member of the SIML. If calling the witness's attention to each face in random order and posing this specific and absolute judgment question is an important factor, such a comparison would be important to make for a more complete study of the factors differentiating SIMLs and SEQLs.

Zimmerman et al. (2006) found that these two variables that have confounded the SIML–SEQL comparison—backloading and number of questions—are critical to sequential superiority. Their research showed that when the traditional SIML and SEQL "packages" were examined (non-backloaded and one question for SIMLs; backloaded and multiple questions for SEQLs), removing the traditional SEQL package also removed sequential superiority compared with traditional SIMLs. Further, no difference in identification rates was found between SIMLs and SEQLs when they both used traditional SIML procedures or when they both used traditional sequential procedures. It may be that SEQL superiority is actually a function of the package of variables that has been uniquely associated with SEQLs. In any case, it seems clear that the value of SEQLs is sensitive to the different variables administered in association with them.

Until these (and perhaps other) factors that differentiate SIMLs and SEQLs are more fully empirically examined, they are confounds preventing clear interpretation of the current body of research. The fact that lineup members are presented one at a time in SEQLs may not in fact be the reason for the sequential superiority effect; it could be due instead to these two other factors of which we are aware and possibly others as well. If we do not understand what the important factors are that differentiate SIMLs from SEQLs, it becomes difficult to know which elements to protect when making recommendations for the practical implementation of SEQLs.

Issues of Data Interpretation

Our review of the SIML–SEQL literature reveals some problems in data interpretation. We discuss analyses concerning PA lineups and false identifications, followed by issues of interpreting multiple identifications from SEQLs.

Use of designated "innocent" suspects. In a typical eyewitness experiment, a stimulus event is followed by a lineup task in which the perpetrator is present or absent. Manipulation of perpetrator presence or absence, and the concomitant variation in identification rates across SIMLs and SEQLs, is central to the claim that SEQLs are superior. However, eyewitness researchers sometimes face a dilemma when they construct PA lineups: They must arbitrarily designate another individual as the "innocent suspect." Many researchers designate someone who is perceived to be physically similar to the perpetrator. Others (e.g., Lindsay, Smith, & Pryke, 1999) have treated each filler as a suspect, calculating lineup measures of functional size and effective size on each configuration, as if the lineups were independent (Malpass et al., in press). Very few researchers have addressed this issue directly, however, and it remains an unresolved matter; no solution we know of is defensible to the exclusion of others. Using a physically similar target substitute, researchers assume (without evidence) that this is what happens in police lineups and, more significantly, they restrict the range of generalization of the experiment to suspects who bear a strong physical resemblance to the perpetrator. Treating all PA lineup members as the suspect may be no better, especially when the procedure is construed as generating multiple lineups and, therefore, as independent tests of the SEQL procedure. We argue that they are not independent tests. An alternative variant of the all-filler lineup strategy may be to introduce corrections into the calculation of the ensuing identifications so as to estimate both "suspect" and "filler" identifications.[5]

However, we are not convinced that this solution is wholly satisfactory either. The suspect in PA lineups assumes a "special status," as fillers are selected on the basis of their similarity to the suspect (notwithstanding the recommendations made by Luus & Wells, 1994). However, we know that choosing fillers, each with reference to physical similarity to the suspect, can result in the suspect having unique properties, such as being the prototype of the group (Laughery et al., 1988; Wogalter et al., 1992). This casts doubt on the idea that the designated suspect and the fillers chosen with respect to him could be suitable fillers for each other.

Interpreting false identifications. The literature reviewed here indicates that PA lineups are constructed using both one-suspect procedures (a perpetrator-substitute and some number of fillers) and all-suspect procedures (no person is designated as a unique substitute for the perpetrator, therefore, they all qualify as perpetrator substitutes). Given these alternatives, data interpretation becomes problematic when results are reported merely as "false identifications" and the

[5]The proportion of correct rejections (pCR) can first be calculated as (number of correct rejections / total number of eyewitnesses). The proportion of suspect identifications can then be estimated as [(1 – pCR) / number of lineup members], and the proportion of filler identifications as [(1 – pCR) – suspect ID proportion]. Consider the situation where 100 eyewitnesses view a six-member perpetrator absent lineup. If 40 eyewitnesses correctly reject the lineup and 60 identify a lineup member, then pCR = 40/100 = .4, Suspect ID proportion = (1 − .40) / 6 = .6/6 = .1, Filler ID proportion = (1 − .40) − .1 = .6 − .1 = .5.

article does not report whether the perpetrator substitute is to be considered a designated suspect, or whether any identification will be considered to be a false (harmful) identification. This is an important point for two reasons. First, an undifferentiated report of false identifications may mean either that witnesses made some percentage of the lineup identifications for a designated suspect or that this percentage of witnesses made an identification of any one of the members of the lineup. Second, the appropriate method of calculating the false identification rate when the perpetrator substitute is not specifically designated as the suspect requires a different procedure. We argue that the term *false identification* is ambiguous if the structure of the PA lineup is not carefully specified and reported. When the procedure used is unclear, and what is known as false identification and filler identification are not distinguished, a statistical summary risks merging them into one index, making cross-study interpretation difficult or misleading. This difficulty is exemplified in Table 1 of Steblay et al. (2001).

Interpreting multiple identifications. Unless witnesses are instructed otherwise, there is a risk that they will identify more than one member from an SEQL. Few researchers have addressed this issue directly, however, and it remains an unresolved matter as no solution we know of is defensible to the exclusion of others. In this literature, multiple identifications have been coded as inaccurate, given that at least one filler is chosen in the combination of responses (Corber, 1995; Lindsay & Wells, 1985; Parker & Ryan, 1993; Rombough, 1994; Smyth, 1994), and have also been coded as nonidentifications, as a witness's credibility is impeached (Lindsay, Pozzulo, Craig, Lee, & Corber, 1997). Some studies exclude these cases from the analysis (Gaitens, Zimmerman, McQuiston, & Malpass, 2002; McQuiston & Malpass, 2002; McQuiston, Salinas, Villegas, Garcia, & Malpass, 2001), whereas others analyze only the first identification in the combination and ignore the other choices made (Sporer, 1993). Similar to our conjecture above, the interpretation of aggregated data based on various methods of analyzing a particular set of identifications may not be straightforward.

Moderators of the Sequential Superiority Effect

In the analysis described below, we attempted to extend the search for moderators of the sequential superiority effect on the basis of many of the variables we have identified above. We followed conventions similar to those reported by Steblay et al. (2001) in their moderator analysis in reporting unweighted average proportions, unweighted effect sizes (r), and Z_{MA} as a test of the difference between SEQL and SIML conditions. In most of the comparisons, we focused on false identifications of lineup members in PA lineups and on correct identifications of perpetrators in PP lineups. We also followed the lead of Steblay et al. in reporting the overall difference between SEQLs and SIMLs. In the last two moderator analyses, we also considered false identifications of designated innocent lineup members.

Similarity of Perpetrator to Suspect

Studies were coded according to the degree of similarity of the perpetrator to the suspect (i.e., the replacement for the perpetrator in PA lineups). Two levels of similarity were generally recognized in articles in the collection, namely, (a) high similarity and (b) low or moderate similarity. Table 3 reports the results of this analysis.

The results suggest that the similarity of suspect to perpetrator does not affect the difference between the SEQLs and SIMLs in the case of false identifications but does affect it in the case of correct identifications from PP lineups. Sequential lineups consistently produce fewer false identifications in PA lineups and produce a similar number of correct identifications to SIMLs when the level of similarity is moderate or low but produce fewer correct identifications when the level of similarity is high.

Backloading of Lineups

Studies were coded according to whether or not backloading was used in SEQLs when making comparisons with SIMLs. Table 4 reports the results of this analysis.

The results of this analysis suggest that backloading does not moderate the difference between SIMLs and SEQLs, except perhaps in the case of false identifications, where SIMLs may result in a larger number of false identifications than SEQLs that are not backloaded. However, we have not tested this contrast formally and base our conclusion only on the apparent difference in effect size.

Stopping Rule

Studies were coded according to whether or not a stopping rule was used when conducting SEQLs. Two stopping rules were identified as being in common use, namely, one in which all the photographs or persons in the lineup were presented to witnesses in SEQLs (i.e., if they chose someone from the lineup, this did not result in the lineup stopping and witnesses were free to choose someone later in the sequence), and an alternative form in which witnesses were instructed to choose only one person from the sequence. Among the study comparisons, 17 allowed witnesses multiple identifications, and 5 study comparisons required that witnesses choose only one lineup member. Table 5 reports the results of this analysis.

These results suggest that the stopping rule does not moderate the difference between SIMLs and SEQLs in the case of false identifications; SEQLs always have fewer associated false identifications. Similarly, the stopping rule does not affect the difference between the lineup types in the case of correct identifications; SIMLs result in more correct identifications. However, what is striking is the disappearance of the overall advantage for SEQLs when considering studies that have used the strictest form of the stopping rule—the difference in favor of

Table 3
Moderator Analysis for Suspect–Perpetrator Similarity Variable

	False ID					Correct ID					Overall correct			
Similarity	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}
High	.27	.47	.22	6.9*	9	.32	.46	.10	3.7*	6	.58	.46	.11	4.5*
Moderate/low	.30	.58	.27	7.1*	15	.46	.52	.06	0.6	13	.63	.44	.19	6.1*

Note. False ID = identification of any member of a perpetrator-absent lineup; correct ID = identification of perpetrator in a perpetrator-present lineup; Seq = sequential lineup; Sim = simultaneous lineup; MA = meta-analysis.
*$p < .05$.

Table 4
Moderator Analysis for Backloading of Lineups Variable

	False ID					Correct ID					Overall correct			
Backloading	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}
Yes	.24	.44	.22	9.11*	18	.36	.52	.15	4.35*	15	.59	.50	.09	5.09*
No	.37	.68	.28	3.92*	7	.33	.46	.14	2.3*	7	.48	.39	.09	1.72*

Note. False ID = identification of any member of a perpetrator-absent lineup; correct ID = identification of perpetrator in a perpetrator-present lineup; Seq = sequential lineup; Sim = simultaneous lineup; MA = meta-analysis.
*$p < .05$.

SEQLs in terms of reduction of false identifications is balanced by the difference in favor of SIMLs in terms of an increase in correct identifications. Some caution may be necessary in interpreting this result, as there were very few studies in this subset (only 5); on the other hand, however, the total number of participants in this subset was over 700, which suggests that it is reliable.

Laboratory Identity

We noted earlier in the article that a large number of the published studies that compare SIMLs and SEQLs emanate from one laboratory, namely, that of R. C. L. Lindsay at Queens University, Canada. It is important to know whether the sequential superiority effect exists for the literature as a whole or whether it is more characteristic of a subset of studies, particularly those that are likely to share possible peculiarities or regularities of laboratory and methodological practice. Studies were therefore coded according to whether they were identified in the Steblay et al. (2001) MA as being authored or coauthored by R. C. L. Lindsay. Table 6 reports the results of this analysis.

The results of this moderator analysis suggest that the pattern of results when comparing SIMLs and SEQLs is different for published studies emanating from the Lindsay laboratory and those emanating from other laboratories. Specifically, in published articles stemming from the Lindsay laboratory, there is an overall advantage for the SEQL (it results in fewer false identifications at a small cost of fewer correct identifications), but in other laboratories there is no overall advantage for the SEQL. In fact, the averaged proportions are equal, with the difference

Table 5
Moderator Analysis for Stopping Rule Variable

	False ID					Correct ID					Overall correct			
Stopping rule	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}
Multiple IDs	.24	.49	.23	8.88*	17	.39	.52	.10	3.71*	14	.64	.49	.15	6.03*
Only one ID	.39	.58	.20	2.50*	5	.35	.49	.15	2.64*	5	.48	.46	.02	0.93

Note. False ID = identification of any member of a perpetrator-absent lineup; correct ID = identification of perpetrator in a perpetrator-present lineup; Seq = sequential lineup; Sim = simultaneous lineup; MA = meta-analysis.
*$p < .05$.

Table 6
Moderator Analysis for Laboratory Identity Variable

	False ID					Correct ID					Overall			
Laboratory	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}
R.C.L. Lindsay	.27	.58	.28	9.6*	10	.54	.62	.08	1.63	6	.67	.47	.20	7.5*
Others	.29	.46	.21	4.8*	14	.28	.46	.17	4.77*	16	.48	.48	.01	0.3

Note. False ID = identification of any member of a perpetrator-absent lineup; correct ID = identification of perpetrator in a perpetrator-present lineup; Seq = sequential lineup; Sim = simultaneous lineup.
$*p < .05$.

between correct identifications (0.18) almost exactly balancing the difference between false identifications (−0.17).

In Table 6 we have treated all identifications of lineup members in PA lineups as false identifications, but one might object that these are not really false identifications, as only identifications of the innocent suspect have any forensic significance or practical implication. When the false identifications are restricted to identifications of the designated innocent suspect, the difference between the SEQLs and SIMLs is large for published studies emanating from the Lindsay laboratory (a difference of 0.25, $r = .31$, $z_{MA} = 8.75$, $p < .05$), and smaller, but significant for other studies (a difference of 0.08, $r = .12$, $z_{MA} = 2.02$, $p < .05$). When correct identifications are considered for this subsample of studies (not all studies used a designated suspect), the published studies from the Lindsay laboratory failed to find a reliable advantage for the SIML over the SEQL (the difference of proportions = .05, $r = .04$, $z_{MA} = .74$, $p > .05$). However, studies that did not emanate from the Lindsay laboratory find a large advantage for SIMLs over SEQLs on correct identifications (difference of proportions = .19, $r = .21$, $z_{MA} = 3.08$, $p < .05$), and this advantage on the correct identifications from PP lineups balances the disadvantage on the false identifications from PA lineups (difference of proportions = .01, $r = .02$, $z_{MA} = .08$, $p > .05$).

Counterbalancing of Lineup Members

The results reported immediately above suggest that studies published by R. C. L. Lindsay differed in some important way from other studies in the meta-analytic set analyzed by Steblay et al. (2001). Upon inspecting our coding of study characteristics (see Table 1), the only difference between published studies coming from that laboratory and other laboratories that appeared to be relatively frequent was the difference in reported counterbalancing of lineup photographs or members. In published studies emanating from Lindsay's laboratory, counterbalancing of stimuli is reported 2 of 11 times, whereas in other studies it is reported 10 of 18 times. We therefore used a coding system for reported counterbalancing as a further moderating variable. Table 7 reports the results of this analysis.

The results of this moderator analysis suggest that counterbalancing (or its absence) strongly influences the difference between SIMLs and SEQLs. Specifically, studies that counterbalance stimuli fail to find an advantage of SEQLs; in fact, such studies reported a marginally significant overall advantage for SIMLs

Table 7
Moderator Analysis for Counterbalance of Lineups Variable

	False ID					Correct ID					Overall			
Counterbalance	Seq	Sim	*R*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}	*N*	Seq	Sim	*r*	Z_{MA}
Yes	.39	.54	.18	2.84*	11	.25	.44	.20	4.16*	14	.38	.43	.05	1.53
No	.24	.54	.28	10.32*	14	.56	.59	.04	0.72	8	.72	.49	.23	9.38*

Note. False ID = identification of any member of a perpetrator-absent lineup; correct ID = identification of perpetrator in a perpetrator-present lineup; Seq = sequential; Sim = simultaneous.
$^{*}p < .05$.

over SEQLs ($Z_{MA} = 1.53$, $p < .06$), that is, just the opposite pattern to the claimed sequential superiority effect. On the other hand, studies that do not counterbalance stimuli reported a heavy overall superiority for SEQLs. Breaking the analysis down into false identifications and correct identifications indicates (a) that counterbalanced studies show a significant advantage for SIMLs in the case of correct identifications, and (b) a similarly significant advantage for SEQLs in the case of false identifications. Studies that do not counterbalance stimuli show (a) an advantage for SEQLs in the case of false identifications but show (b) no difference in the rate of correct identifications.

When false identifications are narrowed down to identifications of a designated suspect, studies that do not report using counterbalancing show a strong advantage for the SEQL over the SIML in that fewer witnesses choose the innocent suspect (difference in proportions = .23, $r = .30$, $z_{MA} = 8.37$, $p < .05$). This same subset of studies shows no advantage for the SIML over the SEQL in correct identifications of the perpetrator in PP lineups (difference in proportions = .05, $r = .05$, $z_{MA} = .62$, $p > .05$). Studies that do report counterbalancing show a significant but reduced advantage for sequential presentation in PA lineups (difference in proportions = .09, $r = .11$, $z_{MA} = 2.12$, $p < .05$) and a strong advantage for simultaneous presentation in PP lineups (difference in proportions = .21, $r = .23$, $z_{MA} = 3.42$, $p < .05$).

Taken as a whole, this moderator analysis shows that the sequential superiority effect may vary as a function of study methodology. With some study designs, this overall effect is apparent, but in others there is no overall advantage of sequential presentation. We have isolated two instances in which the effect is absent. In the first instance, when a strict stopping rule is used with SEQLs (i.e., the lineup is stopped once the witness identifies someone), there is no overall advantage for SEQLs. The benefit of a reduced number of false identifications is balanced by a similar decrease in the number of correct identifications. In the second instance, studies that report counterbalancing lineup order also do not show an overall advantage for SEQLs. Once again, the reduction in false identifications is balanced by a decrease in the number of correct identifications. Indeed, there is some marginal evidence that SIMLs may actually be more accurate than SEQLs when counterbalancing is used. A caveat is in order here, however. Most of the studies in the meta-analysis that were not counterbalanced or that did not report using counterbalancing emanated from one particular laboratory, and it may be some other methodological peculiarity or regularity specific to that laboratory that is responsible for the sequential superiority effect.

Do Absolute Judgments Account for the Sequential Presentation Advantage?

A common and intuitively appealing theoretical explanation for the advantage of SEQLs has to do with the notions of relative and absolute judgment. These terms were first used in the eyewitness literature by Wells (1984) and refer, respectively, to the situation in which witnesses choose the match in the lineup that is closest to their memory of the perpetrator (SIMLs) and to the situation where witnesses choose the closest match to memory but only if it is above some threshold value (SEQLs). The foundational empirical support claimed for the notion of relative judgments was the observation that when the perpetrator is removed from an SIML, witness responses do not gravitate to rejection of the lineup but move instead to the identification of the most frequently chosen filler in the original lineup (Wells, 1984). Recent research by Clark and Davey (2005) replicates Wells's findings and provides evidence of a similar target-to-fillers identification shift for SEQLs (cf. MacLin & Zimmerman, 2003). Researchers have argued in the years since Lindsay and Wells's (1985) original study of relative and absolute judgments that it is this difference that accounts for the SEQL advantage (Dysart & Lindsay, 2001; Kneller et al., 2001; Lindsay & Bellinger, 1999; Lindsay, Lea, Nosworthy, et al., 1991). It has also been suggested that real witnesses be instructed to use absolute judgments when making an identification from a lineup (Kneller et al., 2001). In reviewing the current literature, however, we believe that the notions of relative and absolute judgment may not provide a satisfactory explanation for the SEQL advantage for two reasons. First, there are some important weaknesses in the attempted measurement of these judgment strategies. Second, theoretical accounts rooted in signal detection theory may offer a more parsimonious explanation.

Self-report measures are generally used to assess relative and absolute judgment processes. Using this procedure, participants respond to items addressing the extent to which they compared the photos to each other to narrow the choices (indicating a relative judgment) or compared each photograph separately to their memory image (indicating an absolute judgment; for examples of this, see Dysart & Lindsay, 2001; Kneller et al., 2001; Lindsay & Bellinger, 1999; Lindsay, Lea, Nosworthy, et al., 1991). However, if participants are given six photos to consider and are asked one omnibus question or are presented with six photos and asked an identification question for each one, then posing an additional self-report question as to whether participants compared photos or considered each on its own may add little independent information. Participants can respond to self-report items by deduction from their memory of the task without metacognitive reflection on, or awareness of, their cognitive strategy; indeed, accurate self-reports by witnesses require that they be retrospectively aware of their own judgment strategies, and the likelihood of this level of awareness is questionable (Nisbett & Wilson, 1977). We believe that self-report measures can provide some interesting and important information about what witnesses do in a lineup identification task, but there are limitations in what has been demonstrated thus far in this literature because this method of collecting information has been so heavily relied on. Additionally, we were unable to locate reliability or validity data of the self-report questionnaires currently in use in the literature, therefore, we know little about whether relative and absolute judgments are actually being measured (apart from some degree of face validity) or about whether what is being measured is being measured consistently. Notably, some attempts have been made to implement other measures of judgment processes or to measure them objectively, such as the use of reaction time measures. For example, Sporer (1993) found that, under sequential conditions, correct identifications were made faster than false filler identifications, suggesting that correct choices were more automatic and perhaps linked with absolute processing, whereas false identifications might be associated with relative judgment strategies. Kneller et al. (2001) reported similar findings but also noted that decision times were not correlated with participants' self-reported use of absolute or relative judgment strategies.

The distinction between relative and absolute judgment processes is a good sign of eyewitness researchers thinking theoretically and in terms of cognitive process. However, few connections between this distinction and memory or cognitive theory are made in the eyewitness literature. We believe that it is important to connect eyewitness research and data to the extant body of cognitive and memory theory. There is some sign of this occurring in recent eyewitness research (e.g., Clark, 2003; Gronlund, 2004, 2005). One classic theoretical tradition in memory research is signal detection theory (Green & Swets, 1966; for a review, see MacMillan & Creelman, 1991). Signal detection theory has been used in a range of memory research areas, and there are well-established quantitative implementations of the theory. Signal detection theory was first introduced in the eyewitness literature in 1969 in the context of a study on recognition of own- and other-race faces (Malpass & Kravitz, 1969). However, it has not been used as a theoretical or analytical model in eyewitness research, although many authors have argued for its conceptual suitability (e.g., Malpass & Devine, 1984). Ebbesen and Flowe (2002) have shown how the SEQL and SIML comparison can be modeled within a signal detection theory framework, demonstrating the SEQL advantage to be a function of criterion shift rather than change in discriminability: Eyewitnesses who view SEQLs could simply become more conservative when making identifications, and this explains why they made both fewer positive identifications and fewer false identifications.

From a practical point of view, the reliance in the eyewitness literature on the staged crime experiment has meant that it is not possible to use signal detection theory as a model of eyewitness memory. This is because staged crime experiments yield one data point per participant, and signal detection models require multiple data points. However, Meissner, Tredoux, Parker, and MacLin (2005) have recently developed an experimental procedure and a signal detection model based on dual process theory for application to eyewitness memory. Meissner et al. reported results from four studies that redesigned the classic eyewitness experiment to accommodate empirical signal detection methods, and those results are consistent with the view that SEQLs make witnesses more conservative but not more accurate. In one of the studies reported by Meissner et al., witnesses were given lineup instructions designed to make them very conservative ("only choose if you are 100% certain"), which negated the differences between SIMLs and SEQLs on hit and false positive rates.

This account is different from that given by the relative versus absolute judgment theoretical position. It is too early to say which theoretical account is to be preferred, or whether either accounts for the data more completely than the other. We believe, however, that the use of the signal detection theory model gives researchers access to a paradigm of research that is already established within memory research rather than starting more or less anew. Also, as others have pointed out (e.g., Malpass & Devine, 1984), eyewitness identification tasks are in important respects recognition tasks requiring criterion-related decision making on the part of the witness. It has been shown that response decision processes are important in eyewitness behavior (Steblay, 1997), and signal detection theory has provided a productive theoretical explanation for the analysis of eyewitness choice behavior—quite apart from its contribution over more than five decades to the study and analysis of response decision making in perceptual and memory tasks.

Overall, we believe that as things stand, the current SIML versus SEQL literature provides an underdeveloped theoretical understanding of the decision processes used with respect to lineup presentation methods, and further research aimed at developing a good theoretical understanding of the differences is needed.

Discussion and Conclusion

Our examination of the literature comparing SEQLs and SIMLs has revealed some important limitations as well as some interesting and relevant directions for future research. We have reviewed aspects of the current literature that are methodological, presented the results of a moderator analysis that extends the search for moderators of the SEQL–SIML contrast, and highlighted concerns about the current knowledge and understanding of the theoretical aspects underlying the differences between SEQLs and SIMLs. As members of a scientific community, it seems appropriate to acknowledge the interdependence of theory, methodology, and application in a research field; application requires an understanding of more than the appearance of what works. To gain that understanding requires a theoretical account that has survived the development of rival alternative explanations and competitive empirical evaluations. The Campbellian revolution in social science methodology was about this, if it was about anything (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002; Webb, Campbell, Sechrest, & Schwartz, 1966).

Details of methodology and procedure, as well as the clarity and completeness of reporting, are important aspects of scientific communication and are essential to the interpretation of scientific research. Increasingly, under the criteria of *Daubert* and related U.S. Supreme Court decisions (Studebaker & Goodman-Delahunty, 2002), new forms of, and standards for, scientific evidence are being evaluated in the courts in terms of what good scientific status and practice should involve. Fields of study that do not meet these standards risk failure to achieve recognition. For example, as fingerprint evidence is receiving new scrutiny against scientific standards (e.g., *United States v. Hines,* 1999), and as person identification through DNA evidence has recently gone through the process of approval—through the courts—for widespread use (National Research Council, 1996), it is reasonable to expect that similar evaluation will be given to other new evidential techniques. Experimental psychologists have worked hard in a number of areas to achieve the modest reception currently experienced—which can only be described as sporadic—and it would not further credibility for the field to be weak on some basic scientific attributes, including some highlighted in this article.

The results of our moderator analysis indicate that the SEQL advantage may vary as a function of study methodology in that the effect emerges with some study designs but not with others. Specifically, the stopping rule used when administering SEQLs, and whether counterbalancing of lineup members is used, appears to influence whether the SEQL advantage is found. Another important finding is the contrast between studies authored by R. C. L. Lindsay and those without that authorship, namely, that sequential superiority is found more often in studies produced from that laboratory than from others. We reject any suggestion that there is a bias in favor of SEQLs from that laboratory, particularly because two of the present authors have replicated the overall pattern of the Steblay et al. (2001) MA using the general procedures set forth in publications by Lindsay, and argue that it is variations in methodology that appear to account for whether the SEQL advantage is found. We have indeed found one aspect of methodological practice (i.e., counterbalancing) that seems to shed some light on the laboratory difference. In terms of examining all of the specific methodological aspects that might further distinguish the Lindsay studies from others, we have examined those aspects that we can in this respect, as summarized in Table 1. Of course, a host of estimator and system variables (Wells, 1978) can influence lineup performance, including target exposure (e.g., Laughery, Alexander, & Lane, 1971), stress and intoxication (e.g., Read, Yuille, & Tollestrup, 1992), witness/culprit ethnicity (Meissner & Brigham, 2001), the method used to elicit a description (Sporer, 1996), witnesses' response criterion (e.g., Meissner, 2002), witness expectations (e.g., Douglass & McQuiston-Surrett, in press), and post-identification feedback (e.g., Wells & Bradfield, 1998).[6] Our options remain limited for understanding many elements of laboratory practice implemented, however, because the articles in this literature are often not very detailed on the procedures used (a point described earlier in this article). Overall, the results of this analysis, together with our descriptive analysis, demonstrate the great variability in procedures used when administering SEQLs and the resulting impact on the research findings. This underscores the point of our criticism of the body of work on SIMLs versus SEQLs.

The results of recent field studies of eyewitness identification in several U.S. jurisdictions offer some important findings related to the SIML–SEQL contrast. These studies include one in Hennepin County, Minnesota; a second in Queens County, New York; and a third in three jurisdictions in Illinois (Klobuchar, Steblay, & Caligiuri, in press; Mecklenburg, 2006), all of which were implemented by law enforcement personnel using actual cases. The Illinois study, however, was implemented to make a direct comparison of existing (SIML) practices with SEQL blind protocols advocated as replacements for existing practice. The results from these studies are summarized in Table 8.

The SEQL data from Illinois and Hennepin County are remarkably similar: Approximately 45% of witness decisions were suspect identifications. Illinois and Queens County also yielded similar SIML suspect identifications: 59.9% and 53.6%, respectively. Results showing more suspect identifications from SIMLs than SEQLs are consistent with Steblay et al. (2001). Filler identifications were greater with SEQLs (Hennepin County = 10.3%, Illinois = 9.2%) than with

[6]See Narby, Cutler, and Penrod (1996) and Shapiro and Penrod (1986) for reviews of the impact of system and estimator variables on eyewitness identification performance.

Table 8
Results of Lineup Identifications From Field Studies

Identifications	Sequential procedure: Hennepin County, Minnesota	Sequential procedure: Illinois	Simultaneous procedure: Illinois	Simultaneous procedure: Queens County, New York
Suspect	45.1%	45.0%	59.9%	53.6%
Filler	10.3%	9.2%	2.8%	3.0%
None	44.2%	47.2%	37.6%	43.4%

SIMLs (Illinois = 2.8%, Queens County = 3.0%). Within the Illinois study, SEQLs yielded more nonidentifications than did SIMLs, but the rate of nonidentifications for Hennepin County (SEQL) and Queens County (SIML) were similar. The most facile interpretation here is that sequential superiority is not clearly visible in these findings. However, this is unsatisfying for reasons echoed from the previous discussion. The suspect identifications contain an unknown proportion of false identifications, just as nonidentifications contain incorrect lineup rejections, thus, there lies some uncertainty in drawing conclusions about the data (see Behrman & Davey, 2001). Also, these results do not advance our understanding of the psychological processes involved nor do they point the way toward development of new techniques or procedures for reducing false identifications while maintaining high levels of correct identifications. It is precisely the highly analytical, well-documented, and reported applied research advocated above that will achieve those desired results.

It is common knowledge that eyewitnesses are often mistaken and that their mistakes can have grave consequences. One of the ostensible safeguards against such potential miscarriages of justice is the police lineup, and for over 150 years Anglo-American legal systems have used the same general type of lineup: The suspect is placed among a group of his or her peers who are of sufficient physical resemblance, and the eyewitness is asked to pick the perpetrator. This is certainly an imperfect safeguard, as many have argued, and as multiple recent DNA exoneration cases make clear (see http://www.innocenceproject.org/; Wells et al., 1998). For this and other reasons, it is undoubtedly a good idea to research ways to improve police lineups; eyewitness researchers who have done so have had a palpable impact on public policy concerning eyewitness evidence (Wells et al., 2000). However, it is important to be careful that the pressing needs of public service (in this case, the criminal justice system) do not move us to potential premature advocacy of any line of research and the application of that research. We believe that the recent promotion of the SEQL in research journals, conferences, and policy arenas may be a case in point. On the basis of our analysis, we argue that the research base for SEQLs may not be sufficiently developed from a methodological or theoretical point of view to currently advocate for its implementation to the exclusion of other procedures.

We further argue that a good theoretical understanding of this (and any other) phenomenon is required prior to application—particularly in the implementation of policy, because application occurs in a dynamic environment. Conditions are different from one application venue to another, and compromises from the ideal conditions specified by a set of research results and any formal policy will have to be adapted. The application of SEQLs in a range of states and localities is testament to this. A well worked out theoretical account of the phenomenon is needed in order to adapt practice to achieve the intent of policy in a dynamic application environment. If the research base of policy does not adequately represent the study space of the research domain (the studies implied by the various competing theoretical accounts of the phenomenon), it will be difficult to know what adaptations to develop.

The current review yields two main directions for future research. First, we have outlined a basic question regarding whether the effects of SEQLs have been appropriately examined in the literature. Prior to Zimmerman et al. (2006), the variables uniquely associated with SEQLs (backloading and multiple questions) have consistently confounded the comparison between SIMLs and SEQLs. There are reasons why studies may have found differences between SIMLs and SEQLs that have little to do with whether the presentation is an SIML or SEQL because, along with the change in presentation mode, the SEQL incorporates additional procedural elements that are not included with SIML. Comparisons between these procedures in the literature are confounded by using a combination of certain procedures with SEQLs and a combination of other procedures with SIMLs. For example, in the SEQL package, witnesses should not know the number of lineup members they will see, but this is not the case in the SIML, and this difference is likely to produce a difference in response criterion between witnesses in the respective lineups. A similar difference between the packages called SEQLs and SIMLs occurs in the posing of multiple identification questions to witnesses who view an SEQL and asking witnesses who view an SIML just a single question. This procedure embodies Lindsay and Wells's (1985) idea that SEQLs will reduce comparisons among lineup members and facilitate absolute comparisons for each lineup member with the witness's memory for the perpetrator. This difference can be expected once again to encourage different response criterion levels and is another potential source of confound. An additional confound concerns the overuse of designated suspects in PA lineups who are highly similar in appearance to the perpetrator, which may provide a lure for (false) identification in SIMLs but not in SEQLs. We know very little about the SEQL advantage under PA conditions when the suspect is not highly similar in appearance to the perpetrator, which represents an unknowable proportion of PA lineups in police practice. As it stands, the SIML versus SEQL literature covers only a subset of theoretically interesting and important variables, and studies should be conducted that sample the study space more thoroughly (Malpass, MacLin, Zimmerman, Tredoux, & McQuiston, 2003).

Second, we believe that current explanations for why sequential presentation should reduce both mistaken identifications and correct identifications are underdeveloped. Although the concept of relative judgment is theoretically and intuitively appealing and of considerable interest here, we believe that the way this construct has been tested in much of the literature does not provide a strong basis for concluding that eyewitnesses preferentially use this strategy when choosing from an SIML nor that they preferentially use an absolute judgment strategy when choosing from an SEQL. Because there is a set of variables that differentiates SIMLs and SEQLs, which have not been examined in this literature, as well as alternative theoretical explanations that remain underexplored, there are important questions about where the sequential superiority effect comes from and the specific conditions under which SEQLs are superior (for examples, see Gronlund, 2005; Meissner et al., 2005; Memon & Gabbert, 2003b).

In general, it is important to identify both shortcomings of, and future directions for, a body of literature, and this becomes an even more important consideration when a literature is used as a basis for public policy recommendations. In the early stages of research, questions such as the degree of coverage a line of research has given to its possible study space are perhaps unimportant, especially as the work remains within academia. When the results of a research literature are used to develop and justify public policy recommendations, however, a number of issues become extremely important, such as the size of the literature, the distribution of studies in the study space, potentially confounded comparisons, variations in methodology and analysis, and the lack of a firm theoretical understanding of the reported effects. Research on the difference between SEQLs and SIMLs is very important research—it has strongly influenced a generation of eyewitness researchers and has contributed to the recent collaborations between eyewitness researchers and justice/law enforcement officials (e.g., Technical Working Group for Eyewitness Evidence, 1999). We argue, however, that this particular line of inquiry is in need of amplification. Studies should be conducted that sample the study space more thoroughly, including those that disentangle the differences between SEQL and SIML procedures, so that we can further investigate whether differences in identification accuracy are truly due to mode of presentation. Methodological variations in the administration of SEQLs should continue to be explored. Moreover, greater attention needs to be given to theoretical analyses derived from the study of choice behavior; research should more fully explore alternative theoretical explanations for the differences between SEQLs and SIMLs. As it stands, the existing theoretical formulations do not have strong empirical support, and without an adequate account of why SEQLs should differ from SIMLs, our research cannot move beyond a verificationist epistemology to the hypothetico-deductive standard of empirical sciences.

References

*References marked with an asterisk were included in the studies reviewed in Table 1.

Beaman, A. L. (1991). An empirical comparison of meta-analytic and traditional reviews. *Personality and Social Psychology Bulletin, 17,* 252–257.

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior, 25,* 475–491.

*Bellinger, K. (1995). *Correct lineup rejections as a function of lineup presentation.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

*Blank, H., & Krahe, J. (2000). *The influence of an outstanding similarity between two persons in a lineup on target identification in sequential and simultaneous lineups.* Unpublished manuscript, University of Leipzig, Germany.

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology, 17,* 629–654.

Clark, S. E., & Davey, S. L. (2005). The target-to-foils shift in simultaneous and sequential lineups. *Law and Human Behavior, 29,* 151–172.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally College.

Cooper, H. M., & Hedges, L. V. (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87,* 442–449.

Corber, S. K. (1995). *Identification procedure: Implications for preschoolers' eyewitness identification accuracy.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

*Cox, M. (1992). *The role of estimator and systems variables in face identification.* Unpublished doctoral dissertation, Purdue University.

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

Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

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

Doob, A. N., & Kirshenbaum, H. M. (1973). Bias in police lineups—Partial remembering. *Journal of Police Science and Administration, 18,* 287–293.

*Dormer, G. (1983). *Effect of absolute judgmental processes on eyewitness identification accuracy.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Douglass, A. B., & McQuiston-Surrett, D. (in press). Post-identification feedback: Exploring the effects of sequential photospreads and eyewitnesses' awareness of the identification task. *Applied Cognitive Psychology.*

*Dysart, J. E. (1999). *The effects of delay on eyewitness identification accuracy.* Unpublished master's thesis, Queen's University, Kingston, Ontario, Canada.

*Dysart, J. E., & Lindsay, R. C. L. (2001). A preidentification questioning effect: Serendipitously increasing correct rejection. *Law and Human Behavior, 25,* 155–165.

Ebbesen, E. B., & Flowe, H. (2002). *Simultaneous v. sequential lineups: What do we really know?* Unpublished manuscript.

*Gaitens, K., Zimmerman, L., McQuiston, D. E., & Malpass, R. S. (2002, March). *Structural bias in eyewitness lineups: When does the suspect stand out?* Poster session presented at the annual meeting of the American Psychology-Law Society, Austin, TX.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5,* 3–8.

Gonzalez, R., Davis, J., & Ellsworth, P. C. (1995). Who should stand next to the suspect? Problems in the assessment of lineup fairness. *Journal of Applied Psychology, 80,* 525–531.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gronlund, S. D. (2004). Sequential lineups: Shift in criterion or decision strategy? *Journal of Applied Psychology, 89,* 362–368.

Gronlund, S. D. (2005). Sequential lineup advantage: Contributions of distinctiveness and recollection. *Applied Cognitive Psychology, 19,* 23–37.

*Hannaford, K. M. (1985). *Relative versus absolute judgments in lineup identifications.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Haw, R. M., & Fisher, R. P. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *Journal of Applied Psychology, 89,* 1106–1112.

*Jacob, P. (1994). *The feasibility of using multiple culprit sequential lineups.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Kassin, S., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56,* 405–416.

Klobuchar, A., Steblay, N., & Caligiuri, H. (in press). Improving eyewitness identifications: Hennepin County's blind sequential lineup pilot project. *The Cardozo Public Law, Policy and Ethics Journal.*

*Kneller, W., Memon, A., & Stevenage, S. (2001). Simultaneous and sequential lineups: Decision processes of accurate and inaccurate eyewitnesses. *Applied Cognitive Psychology, 15,* 659–671.

Kolata, G., & Peterson, I. (2001, July 21). New Jersey is trying new way for witnesses to say, "It's him". *The New York Times.*

*Laldin, S. (1997). *Contextual effects on lineup identification of multiple culprits.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Laughery, K. R., Alexander, J. F., & Lane, A. B. (1971). Recognition of human faces: Effects of target exposure time, target position, and type of photograph. *Journal of Applied Psychology, 59,* 490–496.

Laughery, K. R., Jensen, D. G., & Wogalter, M. S. (1988). Response bias with prototypic faces. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (pp. 157–162). Oxford, England: Wiley.

Levi, A. M., & Lindsay, R. C. L. (2001). Lineup and photospread procedures: Issues concerning policy recommendations. *Psychology, Public Policy, and Law, 7,* 776–790.

Lindsay, R. C., & Bellinger. K. (1999). Alternatives to the sequential lineup: The importance of controlling the pictures. *Journal of Applied Psychology, 84,* 315–321.

*Lindsay, R. C., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76,* 741–745.

*Lindsay, R. C., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76,* 796–802.

*Lindsay, R. C., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior, 18,* 527–541.

Lindsay, R. C., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21,* 391–404.

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

Lindsay, R. C. L. (1999). Applying applied research: Selling the sequential line-up. *Applied Cognitive Psychology, 13,* 219–225.

Lindsay, R. C. L. (2002). *Limitations of sequential lineups.* Unpublished manuscript.

Lindsay, R. C. L., Smith, S. M., & Pryke, S. (1999). Measures of lineup fairness: Do they postdict identification accuracy? *Applied Cognitive Psychology, 13,* S93–S107.

Lindsay, R. C. L., & Turtle, J. (2001). *Best practices for suspect identification: 'Blind' sequential lineup procedures.* Unpublished manuscript.

Luus, C. A., & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology, 79,* 714–723.

MacLin, O. H. (2001). *Towards an ecological approach in examining the diagnosticity of sequential and simultaneous lineup presentations.* Unpublished manuscript.

MacLin, O. H., & Zimmerman, L. A. (2003, July). Evaluating relative vs. absolute judgments when identifying members of another race. In C. A. Meissner (Chair), *Cross-race effect in memory for faces: Developmental, theoretical and applied implications.* Symposium presented at the meeting of the Society for Applied Research in Memory & Cognition, Aberdeen, Scotland.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5,* 299–309.

Malpass, R. S., & Devine, P. G. (1983). Measuring the fairness of eyewitness identification lineups. In S. Lloyd-Bostock, & B. Clifford (Eds.), *Evaluating witness evidence* (pp. 81–102). London: Wiley.

Malpass, R. S., & Devine, P. G. (1984). Research on suggestion in lineups and photospreads. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 64–91). Melbourne, Australia: Cambridge University Press.

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other "race." *Journal of Personality and Social Psychology, 13,* 330–334.

Malpass, R. S., MacLin, O. H., Zimmerman, L. A., Tredoux, C. G., & McQuiston, D. E. (2003, July). A study space analysis for policy development: Application to simultaneous vs. sequential lineups. In R. S. Malpass (Chair), *Simultaneous or sequential lineups? New data, new perspectives, new looks at old questions, and a study space based policy analysis.* Symposium presented at the meeting of the American Psychology-Law Society, Edinburgh, Scotland.

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (in press). Lineup construction and lineup fairness. In R. Lindsay, D. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology: Volume 2. Memory for people.* Mahwah, NJ: Erlbaum.

*Martins, S. (1996). *The effects of changes in the appearance and lineup position of targets on eyewitness identification.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Marwitz, D. B., & Wogalter, M. S. (1988). *Bias in photo-spreads of faces: A comparison of two lineup construction methods.* Paper presented at the annual proceedings of the Human Factors Society, Santa Monica, CA.

McQuiston, D. E., & Malpass, R. S. (2002). Validity of the mock witness paradigm: Testing the assumptions. *Law and Human Behavior, 26,* 439–453.

*McQuiston, D. E., Salinas, R., Villegas, E., Garcia, G., & Malpass, R. S. (2001, May). *Utility of the sequential lineup procedure: Exploring generalizability issues.* Poster session presented at the annual meeting of the Western Psychological Association, Maui, HI.

Mecklenburg, S. H. (2006, March). *Report to the legislature of the State of Illinois: The Illinois pilot program on sequential double-blind identification procedures.* Springfield, IL: Illinois State Police.

*Meissner, C. A. (2002). Applied aspects of the instructional bias effect in verbal overshadowing. *Applied Cognitive Psychology, 16,* 911–928.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7,* 3–35.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33,* 783–792.

*Melara, R. D., DeWitt-Rickards, T. S., & O'Brien, T. P. (1989). Enhancing lineup identification accuracy: Two codes are better than one. *Journal of Applied Psychology, 74,* 706–713.

*Memon, A., & Bartlett, J. (2002). The effects of verbalization on face recognition in younger and older adults. *Applied Cognitive Psychology, 16,* 635–650.

*Memon, A., & Gabbert, F. (2003a). Improving the identification accuracy of senior witnesses: Do pre-lineup questions and sequential testing help? *Journal of Applied Psychology, 88,* 341–347.

*Memon, A., & Gabbert, F. (2003b). Unraveling the effects of sequential presentation in culprit present lineups. *Applied Cognitive Psychology, 17,* 703–714.

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

National Research Council. (1996). *Evaluation of forensic DNA evidence: Update on evaluating DNA evidence.* Washington, DC: National Academies Press.

*Newman, K. (1998). *The effects of a weak memory trace on sequential and simultaneous lineup identifications.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Nisbett, R. E., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental process. *Psychological Review, 84,* 231–258.

*Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior, 17,* 11–26.

*Parker, J., Tredoux, C., & Nunez, D. (2000, March). *Lineup measures, lineup procedure, and optimality of encoding.* Poster session presented at the annual meeting of the American Psychology-Law Society, New Orleans, LA.

*Phillips, M. R., McAuliff, B. D., Kovera, M. B., & Cutler, B. L. (1999). Double-blind photoarray administration as a safeguard against investigator bias. *Journal of Applied Psychology, 84,* 940–951.

Read, J. D., Yuille, J. C., & Tollestrup, P. (1992). Recollections of a robbery: Effects of arousal and alcohol upon recall and person identification. *Law and Human Behavior, 16,* 425–446.

*Rombough, V. J. (1994). *The effects of lineup presentation and practice on lineup identification accuracy in child eyewitnesses.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

*Shapiro, L. R., & Hiatt, T. A. (2003). *How to pick a "perp": An investigation of photographic lineup procedures.* Unpublished manuscript.

Shapiro, P., & Penrod, S. D. (1986). A meta-analysis of the facial identification literature. *Psychological Bulletin, 100,* 139–156.

*Smyth, L. (1994). *Sequential presentation and practice: The jury is still out.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

*Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78,* 22–33.

Sporer, S. L. (1996). Psychological aspects of person descriptions. In S. L. Sporer, R. S. Malpass, & G. Köhnken (Eds.), *Psychological issues in eyewitness identification* (pp. 53–86). Mahwah, NJ: Erlbaum.

Steblay, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21,* 283–297.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 25,* 459–473.

Studebaker, C. L., & Goodman-Delahunty, J. (Eds.) (2002). Expert testimony in the courts: The influence of the *Daubert, Joiner,* and *Kumho* Decisions, Part 2 [Special issue]. *Psychology, Public Policy, and Law, 8*(3).

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement* (NCJ No. 178240). Washington, DC: National Institute of Justice.

Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior, 22,* 217–237.

Tredoux, C. G. (1999). Statistical considerations when determining measures of lineup size and lineup bias. *Applied Cognitive Psychology, 13,* S9–S26.

Tunnicliff, J. L., & Clark, S. E. (2000). Selecting foils for identification lineups: Matching suspects or descriptions? *Law and Human Behavior, 24,* 231–258.

United States v. Hines, 55 F. Supp. 2d 62 (D. Mass. 1999).

*Vanderwal, A. (1996). *The effects of the sequential presentation of lineups with multiple culprits on eyewitness identification.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

*Varrette, T. K. (1994). *The effects of similarity-to-suspect versus match-to-description strategies for selecting lineup foils when culprit appearance has been altered.* Unpublished undergraduate thesis, Queen's University, Kingston, Ontario, Canada.

Webb, E. J., Campbell, D. T., Sechrest, L., & Schwartz, R. D. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36,* 1546–1557.

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14,* 89–103.

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48,* 553–571.

Wells, G. L. (2002, Fall). Expert opinion. *American Psychology-Law Society News, 22,* pp. 14–15.

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect": Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83,* 360–376.

Wells, G. L., & Luus, C. A. E. (1990). Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personality and Social Psychology Bulletin, 16,* 106–117.

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55,* 581–598.

Wells, G. L., Seelau, E. P., Rydell, S. M., & Luus, C. A. E. (1994). Recommendations for properly conducted lineup identification tasks. In D. F. Ross, J. D. Read, & M. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 223–244). New York: Cambridge University Press.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 23,* 603–647.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115–1125.

*Willis, C. E. (1990). *Strategies in eyewitness identification.* Unpublished doctoral dissertation, University of Kansas.

Wogalter, M. S., Malpass, R. S., & McQuiston, D. E. (2004). A national survey of U.S. police on preparation and conduct of identification lineups. *Psychology, Crime & Law, 10,* 69–82.

Wogalter, M. S., Marwitz, D. B., & Leonard, D. C. (1992). Suggestiveness in photospread lineups: Similarity induces distinctiveness. *Applied Cognitive Psychology, 6,* 443–453.

Yarmey, A. D., & Morris, S. (1998). The effects of discussion on eyewitness memory. *Journal of Applied Social Psychology, 28,* 1637–1638.

Zimmerman, L. A., Malpass, R. S., & MacLin, O. H. (2006). *Unconfounding the simultaneous vs. sequential lineup comparison.* Manuscript submitted for publication.

Received December 20, 2005
Revision received March 18, 2006
Accepted March 19, 2006 ■