



Public policy and sequential lineups

Roy S. Malpass^{1*}, Colin G. Tredoux² and Dawn McQuiston-Surrett³

¹University of Texas at El Paso, El Paso, TX, USA

²University of Cape Town, Rondebosch, SA

³Arizona State University, Glendale, AZ, USA

The claim that sequential lineups are superior to simultaneous lineups and that our knowledge of sequential lineups is sufficient to warrant their being required by law is reviewed for the validity of both strong and weak claims of sequential superiority, adherence to principles of research design, and the needs of public policy. We conclude, (1) there is little evidence to support the claim that sequential presentation of photos is responsible for lower levels of false identifications, (2) the evidence is weak that the aggregation of factors commonly labeled as the sequential lineup together produce lower levels of false identifications without additional offsetting effects, (3) much of the literature contains several confounds in research design and additional offsetting effects that question its overall utility, (4) recent research shows that the superiority of sequential lineups is restricted to specific ranges on other study design variables, and (5) the corpus of research on sequential lineups does not satisfy the needs of policy sufficiently to justify its mandated use as the required identification procedure throughout the criminal justice system.

It is likely that eyewitness researchers working in the last half of the 20th century knew that ultimately their research was about policy, but until recently this implication was not well appreciated. Policy reform of eyewitness identification procedures is important, and many aspects of eyewitness identification procedures are in need of revision (Wells *et al.*, 1998). Some reforms have been adopted by law enforcement, with changes to the instructions used when administering lineups being perhaps the most common. Important reforms are advocated by the National Institute of Justice Technical Working Group on Eyewitness Evidence (1999) and non-governmental organizations interested in justice issues (Innocence Project, Justice Project).

Academic researchers generally do not encounter the complexities of policy development. Ordinarily they are free to follow their own lines of inquiry, to devise new concepts, measurements and forms of application, without responsibility to identify and

*Correspondence should be addressed to Dr Roy S. Malpass, Department of Psychology, University of Texas at El Paso, 500 W. University Avenue, El Paso, TX 79968, USA (e-mail: rmalpass@utep.edu).

explore directions not taken or questions that are unaddressed. Cleaning up after oneself is a far less interesting way to spend valuable time and resources than exercising continued vision and discovery. In many ways the freedom of scientific investigation has served us very well. Balance and competitive evaluation of one line of research against others is something done by a community of scientists, outside of the linear development of a particular idea within a particular laboratory, or by a particular scientist. Winnowing the various attempts to understand a given phenomenon, and delineating their limitations, often falls to other researchers.

In contrast, public policy formation often needs to be aware of what has not been studied and the limiting conditions and assumptions in a particular research context (Malpass *et al.* 2008). New roles and evaluative criteria are thrust upon scientists when developing knowledge that is no longer merely academic, but which intends to change the ways things are done in public service and by public agencies.

There are no uniform criteria across governments or agencies for evaluating proposed reforms, the adequacy of competing proposals, and in the case of scientific domains, the coverage, stability and validity of the science underlying them. There is no equivalent of *Consumer Reports* for performing utilitarian policy evaluations. A cost/benefit (utility) analysis of the options is in order (Malpass, 2006a). Many decision strategies can be used for evaluating proposed reforms. But generally, the greater the potential costs, the greater the care needed in identifying and evaluating alternatives. Policy development should be viewed as the creation of a public good, and this carries with it great responsibility, shared by those who attempt to use their work to change policy.

Desiderata for a research domain serving as a policy base

The needs of a research domain as a policy base include the following (discussed in more detail elsewhere, see Malpass, 2006a; Malpass *et al.*, 2008; McQuiston-Surrett, Malpass & Tredoux, 2006).

A corpus of published research, peer reviewed, and in the public domain, that has:

- (1) Consistent findings, as revealed by replication and meta-analysis, across a diversity of methods, materials and research populations,
- (2) Studied an appropriate range of important variables related to the environment of application and surrounding the policy solution being developed,
- (3) An absence of important and systematic methodological errors and omissions, including confounded comparisons, and
- (4) Covered important policy alternatives.

The reform most aggressively advocated in the eyewitness area has been the replacement of the most widely used identification procedure – simultaneous presentation of photographs (Wogalter, Malpass & McQuiston, 2004) – with a new procedure – sequential presentation of photographs. As much as reform is needed, we argue there are some important difficulties with the empirical and theoretical aspects of the literature on sequential lineups that makes heavy marketing inadvisable at best, and at worst destructive of psychology's ability to continue as a source of technological development for law enforcement. We think it is not a reliable solution to the problems of eyewitness identification, and under some circumstances it may be worse than the

present procedure (Carlson, Gronlund, & Clark, 2008; Clark & Davey, 2005; Clark, Howell & Davey, 2008; Gronlund, 2005; Malpass, 2006a). Other aspects of the eyewitness identification process, e.g. instructions to witnesses (Clark, 2005), and lineup construction (Clark, Howell, & Davey, 2007) – need attention before the applied effectiveness of sequential lineups can be evaluated.

Our concern about the state of the science regarding the relative merits of sequential and simultaneous lineups stems from the systematic nationwide advocacy of sequential lineups as a wholesale replacement for simultaneous lineups as a matter of policy or law. Rather than taking an advocacy position for simultaneous lineups and against sequential lineups, our thinking is quite different. Ross and Malpass (2008) put it this way:

Our primary concern is that psychological scientists will have to take back the recommendation that sequential lineups will solve—or even go a long way to diminish—the problems of false identification and therefore wrongful conviction. Questions about the adequacy of a research domain for supporting sweeping policy changes have not been resolved, and criteria for making such judgments about domain adequacy remain largely undeveloped (but see Malpass, MacLin, Zimmerman, Tredoux, & McQuiston, 2003)¹. Our focus on the strength of both the theoretical and empirical base (McQuiston-Surrett et al., 2006) are well known, as is our analysis of the utility of simultaneous and sequential lineups as policy alternatives and the conditions under which this might change (Malpass, 2006a, 2006b, p. 16). While sequential advocates favor a particular family of lineup procedures, we favor a broader search for ways to confront identification errors—both failures to identify offenders and failures to reject identification of innocent suspects—that are theoretically well understood and empirically stable.

A series of considerations raise concerns for the authoritativeness of scientifically based policy developments in our field and have implications for the future of such endeavours, including: the strength of the empirical aspects of the literature regarding the comparison; the breadth of development of other available policy alternatives for lineup administration; considerations surrounding the evaluation of any policy proposals; and the unfortunate spectacle of scientists becoming embroiled in centuries-old adversarial wrangling in legal circles.

When one will force public life to be conducted in a particular way, according to certain principles and procedures, there is a responsibility to be sure that the principles have merit and substance and the procedures have the effects claimed. We do not believe that the program of research required to reach such conclusions has been carried out in the case of sequential lineup advocacy. And for this reason, we believe that the advocacy for systematically replacing other lineup forms with an undefined form of sequential administration is not merely premature, but ill advised.

Strong and weak forms of the sequential superiority claim

At its simplest, the claim of sequential superiority is that presenting lineup members sequentially reduces the rate of (false) identifications in perpetrator absent lineups to a level lower than obtained when presenting lineup members simultaneously. A corollary claim is that this difference is not offset by a concomitant decrease in correct identifications from perpetrator present lineups.

¹ More recently published as Malpass et al. (2008).

The original Lindsay and Wells (1985) paper, and much of the work that has followed, uses the qualifying word 'sequential' in the phrase 'sequential lineup' to suggest that there is something significant and important *per se* about presenting lineup members to witnesses one at a time. Wells (1993) leaves no doubt about this:

A sequential lineup is a procedure in which each lineup member is presented to the eyewitness individually in sequence. The eyewitness must decide at the time of each initial presentation whether that lineup member is the culprit. A simultaneous lineup, which is traditionally used by police, presents the eyewitnesses with all lineup members at once. (Wells, 1993, p. 561).

The theoretical explanation offered in the 1985 paper - and in many subsequent papers - for sequential superiority, is that sequential presentation reduces the tendency of witnesses to use a 'relative judgment strategy' (this argument is commonly attributed to Wells, 1984, but a very similar argument was made over 20 years earlier by Williams and Hammelman, 1963). Exposing witnesses to lineup members one at a time reduces the ability of witnesses to compare lineup members to each other, and to select that person amongst them who most resembles the witness' memory of the perpetrator. Instead, each lineup member is presumed to be compared to some absolute standard of similarity or resemblance to the original perpetrator. Whereas sequential lineups promote absolute judgments, simultaneous lineups are postulated to promote relative judgments:

. . . sequential presentation of lineups can reduce false identifications of innocent suspects by reducing eyewitnesses' reliance on relative judgment processes (Lindsay and Wells, 1985, p. 556)

It is possible to distinguish strong and weak forms of the claim that sequential lineups are superior to simultaneous lineups. The *strong* form is that sequential presentation itself is the source of, and necessary condition for the superiority. The *weak* form is that a sequential lineup 'package' is effective, in which (i) a number of lineup members are presented one at a time (but witnesses do not know how many are in the sequence); (ii) witnesses must make a continuing 'no' or terminal 'yes' decision for each person presented to them, and (iii) witnesses may not see any lineup member more than once. The suspect should not be presented first in the sequence, and his/her position should preferably be randomly determined. This combination will produce a decreased rate of false positive identifications, which will not be offset by an equivalent decrease in the rate of correct positive identifications. In the weak form of the claim, it is not of interest whether the sequential presentation of the lineup is necessary to its apparent advantage. That the combination produces a more favourable ratio of false positive identifications to correct positive identifications is all that matters.

A careful re-reading of Lindsay and Wells (1985, p. 559) suggests that the researchers combined manipulations to achieve a strong reduction of the tendency of eyewitnesses to make relative judgments. Sequential presentation was offered in several places as the key manipulation, but as it was confounded with other manipulations, there is no direct evidence that sequential presentation in and of itself had any effect on identification rates. Here is an excerpt from Lindsay and Wells (1985):

Because the idea is to prevent as much as possible the opportunity for eyewitnesses to make relative judgments . . . the eyewitness sees one lineup member at a time, makes a yes/no decision for each lineup member on first viewing, and is not informed of the total number to be viewed (p. 559)

What evidence is in favour of the strong and weak forms of the claim that sequential lineups produce more diagnostic identifications (i.e. the ratio of hits to false positives is

greater than for simultaneous lineups)? There is no clear evidence we know of that demonstrates that sequential presentation is necessary to the identification advantage reported for sequential lineups. There simply is no corpus of studies that holds constant the other manipulations introduced alongside sequential presentation in the original procedure. A few studies have investigated alternate accompanying manipulations to sequential presentation, but it is doubtful that they attest in any way to the effectiveness of sequential presentation *per se*.

Lindsay, Lea and Fulford (1991), for instance, compared sequential lineups in which the witnesses either knew or did not know the size of the lineup. They reported an advantage for the latter condition, but this tells us nothing about the effectiveness of sequential presentation, *per se*. They report comparisons to a 'simultaneous lineup condition', and note that both versions of the sequential lineup produced superior identification results when compared to this baseline. It is significant that they did not manipulate any aspect of the simultaneous lineup, thus missing an opportunity to disentangle the effects of sequential presentation and the ancillary manipulations that confound our understanding of its effect.

This confound i.e. of sequential presentation and the ancillary manipulations of the Lindsay and Wells sequential lineup, is, of course, not intractable. The unique effect of sequential presentation can be studied experimentally, that is by the classic controlled manipulation and measurement of variables that is canonical to empirical psychology. *The point is that this has not been done, despite it being well within our research potential.* As a result, we do not know whether it is the sequential presentation itself that produces the apparent identification advantage, or whether the same advantage would be achieved if we harnessed the ancillary manipulations to the traditional simultaneous lineup.

Because of the confound it is misleading to claim that sequential lineups are better than the traditional simultaneous lineups used by police forces all over the world (since at least the 1860s). Nothing much is known about the effects of sequential presentation *per se*. Indeed, we believe that the apparent ingenuity of sequential presentation has been used to persuade lawmakers of the value of the sequential lineup procedure. Reports in national newspapers about sequential lineups are indicative. Consider this report, from a 2001 edition of the New York Times:

Professor Lindsay would stage a mock crime . . . [and] then show the witnesses a traditional lineup of suspects . . . About 20% to 40% of the witnesses mistakenly identified someone as the criminal. When the same suspects were put in a sequential lineup, and the eyewitnesses were shown photographs one at a time, and only once, the rate of false identifications dropped to less than 10%. . . . The reason that sequential lineups work is rather simple. In simultaneous lineups, Professor Lindsay said, witnesses are able to compare individuals, choosing one from the group who looks the most like the person they think they saw commit the crime. But a sequential lineup limits the ability to compare. (retrieved from <http://query.nytimes.com/gst/fullpage.html?res=9F05E2D7123AF932A15754C0A9679C8B63&sec=&spon=&pagewanted=2> on 30 January 2008)

In this account, the simple but rather ingenious innovation of presenting people one at a time produces as few as one quarter of the number of false identifications of traditional police methods. And the reason for this success is that when lineup members are presented sequentially *per se*, a witness is unable to compare them against each other. This claim of dubious validity is thus broadcast nationally and support is drummed up for a particular legal reform.

We do not dismiss the possibility that the sequential presentation of lineup members may be the active ingredient in the Lindsay-Wells lineup. Without it, the false positives might well not decline as precipitously as claimed above, and in the Steblay, Dysart, Fulero and Lindsay (2001) meta-analysis. Indeed, presenting lineup members one at a time strikes us as smart, and exactly the kind of idea that merits careful, scientific investigation. This has not happened. The focus, instead, has been on showing that a bundle of practices (the Lindsay-Wells lineup) provides protection against some of the mishaps of police lineup practice (e.g. poorly matched foils, clothing bias), and on punting this bundle as a wholesale replacement for extant procedures.

The strong form of the claim that sequential lineups have an identification advantage over simultaneous lineups is not supported by the literature. We do not know what presenting lineup members one at a time leads to. It would be useful to know.

Some recognize the lack of support for the strong form of the claim, but argue that in applied psychology and law a weaker claim may be sufficient. Lindsay goes further to claim that the original intention was not premised on sequential presentation, but on a combination of features. The relevant question, he has argued, is whether the combination of features in the Lindsay-Wells lineup is associated with better identification outcomes. If such a combination yields better results than the traditional, simultaneous lineup currently used in police practice, then by the principle of 'best practices', we should use the Lindsay-Wells lineup (Levi & Lindsay, 2001). If this is the case, should we be concerned whether it is the sequential presentation that does the trick? It does not matter, the argument goes, whether it is the sequential presentation, or the terminating affirmative, or the concealment of nominal size, or any other feature, as long as the whole bundle is reproduced.

Does it matter that one has to fall back on weaker claims than those originally postulated? We think it does in the case of the claim of sequential superiority. The arguments for the creation of the Lindsay-Wells lineup (Lindsay & Wells, 1985) followed from an earlier paper by Wells (1984), and were premised on the notion of a relative judgment strategy. The lynchpin in this argument is the notion that presenting lineup members sequentially reduces the tendency of witness to arrive at an identification decision on the basis of comparison of lineup members to each other, rather than to their memory of the offender. To argue that it does not matter what aspect of the Lindsay-Wells lineup is effective, or whether sequential presentation has any role in the overall effect, is to discard an important part of the intellectual provenance of the sequential lineup. It also limits the development of a specific application. If the sequential lineup package were to produce results superior to other forms of identification procedure because it asks a specific identification of each face, or because it creates a higher decision criterion for the witness - rather than because the faces are presented sequentially, then the modifications of current practice needed to achieve improved results would not include sequential presentation, and the new procedures would be quite different, targeted at the portion of the package that is the active ingredient.

Despite this intellectual compromise, we may still ask what empirical research tells us about the weak form of the claim that sequential lineups are better than simultaneous lineups. There has been a fair amount of research on the weak form of the claim, starting with the original paper by Lindsay and Wells (1985), and followed by over 50 subsequent studies, including one published (Steblay, Dysart, Fulero, & Lindsay, 2001) and one unpublished (Steblay, 2007) meta-analysis.

The meta-analysis

The social and health sciences are particularly prone to a range of sampling and measurement errors, and empirical findings in these disciplines consequently show considerable variation across studies. For instance, some studies find that use of the Lindsay-Wells lineup leads to a lower false positive rate, but other studies do not. Statistical solutions to this problem evolved over a number of decades, and are now known collectively as ‘meta-analysis’ (e.g. Hunter & Schmidt, 1990; Rosenthal, 1991). Disagreements in the literature are assessed by aggregating individual study effect sizes, correcting for sampling, measurement, and other forms of artefactual variation, and using heuristic aids to decide what constitutes systematic variation in effect sizes. Meta-analysis has by and large usurped the traditional literature review of empirical psychology, and threatens in the view of some to also usurp the important task of methodological evaluation of research literatures. As all psychologists trained since the 1960s will know, Donald Campbell and his colleagues (e.g. Campbell & Stanley, 1966; Cook & Campbell, 1979) developed a sophisticated way of thinking about psychological knowledge, and of evaluating research studies. Their method of assessing the internal and external validity of studies is taught in almost all undergraduate and graduate courses on research methods, and is one of the canons of research psychology, worldwide. Meta-analysis proposes a different way of resolving methodological questions, which is to keep an ‘open mind’ about whether the methods used in a study compromise its findings, rather than applying Campbell’s razor. A statistical moderator analysis is then conducted on the corpus of studies, and claims about methodological flaws are tested in accordance with whether they emerge in the analysis as moderators. This technique has merit, and we have used it ourselves in a re-meta-analysis of Steblay *et al.* (2001) (McQuiston-Surrett *et al.*, 2006). However, there are many methodological aspects of a research tradition about which a meta-analysis must necessarily be mute. Campbell’s razor, on the other hand, is not mute to these aspects, and when applied to the extant literature on sequential lineups has shown that there are significant shortcomings with what one might call the basic Lindsay-Wells lineup study.

The meta-analysis by Steblay *et al.* (2001) coded for a limited number of potential moderator variables. From our theoretical and methodological analysis of the literature, we identified a number of further, potentially interesting moderators. The first concerned the similarity of lineup members, particularly the similarity of suspect to foils, and suspect to perpetrator. Many studies did not report this (as we believe they ought), but we noticed in our reading of the literature that those that did tended to use suspects that were very similar to the perpetrator (i.e. ‘dead ringers’). This manipulation may affect the external validity of studies of the Lindsay-Wells sequential lineup – it is unlikely that the perpetrator and innocent suspect in a real lineup will resemble each other strongly. In our moderator analysis, we did indeed find an effect for similarity of perpetrator and suspect – in particular, when suspect and perpetrator were highly similar, Lindsay-Wells sequential lineups yielded fewer correct positive identifications than simultaneous lineups. This result accords with the argument first made by Ebbesen and Flowe (2002), that sequential superiority may simply be a function of a change in response criterion, rather than discrimination accuracy – witnesses in the Lindsay-Wells sequential lineup become more cautious about making identifications, both in perpetrator present and perpetrator absent lineups. Empirical evidence in favour of this claim is also reported by Meissner, Tredoux, Parker and MacLin (2005). But the most significant finding from our re-meta-analysis is the dependence of sequential superiority on which

laboratory experimental studies were conducted in, and on whether counterbalancing was used.

Knowing, as any experimenter does, that the subject population available to any particular university laboratory is likely to be regional in character, and that the same experimental materials (e.g. face image stimuli, distracter tasks, simulated crime movies) tend to be used many times in any particular laboratory, we explored in the moderator analysis whether sequential superiority is specific to a particular laboratory, or represents the population of studies. Our moderator analysis showed that, indeed, the sequential superiority effect appears to be specific to the Lindsay laboratory at Queens University, Canada. The studies emanating from other laboratories report, in aggregate, a decrease in false positive identifications when the Lindsay-Wells sequential lineup is used, but also report an equivalent decrease in correct positive identifications. The studies on which R. C. L. Lindsay is listed as an author, on the other hand, report a decrease in false positive identifications that is greater than the decrease in correct positive identifications.

This moderator finding seemed like an ‘empty moderator’ at first, as it didn’t suggest a clear reason why the sequential superiority effect should be specific to a single laboratory. On a re-inspection of our methodological review of the literature, we noticed that most of the studies from the Lindsay laboratory did not report counterbalancing the order of the perpetrator/suspect in their lineups, as did a number of the studies in the literature not emanating from the Lindsay laboratory. We therefore decided to conduct a moderator analysis, using reported counterbalancing/failure to report counterbalancing as a binary moderator. As we report in McQuiston-Surrett *et al.* (2006), studies that report counterbalancing do not, on aggregate, show the sequential superiority effect, whereas studies that fail to report counterbalancing do show the sequential superiority effect.

We conclude the evidence for the weak form of the claim that sequential lineups are superior is limited. There is some evidence for the weak form of the claim, but only for studies that fail to report (and presumably implement) counterbalancing the position of the suspect/perpetrator in the lineup (McQuiston-Surrett *et al.*, 2006).

Is the evidence sufficient to support policy recommendations?

Investigation of sequential lineups is rife with inadequate science. Most important is the confounded comparison between simultaneous and sequential lineups. Aggregating multiple components and then calling the aggregate superior to some other aggregate without seeking analysis of just which part of the aggregation is the active ingredient serves neither the purposes of knowledge nor application. Both require that researchers understand how the components work together, if one or the other is a necessary condition for the operation of a third, the rank order of their independent contributions, which components are more sensitive in the application environment (e.g. to training needs, to consistency of application) and which are the most costly to implement. Existing research has done none of these things: it failed to do the one thing that true experiments are best at doing – isolating variables for manipulation and studying them independently. Additional methodological shortcomings are discussed by Malpass, *et al.* (2008) and McQuiston-Surrett *et al.* (2006). This is no exercise in methodological purity. All have implications for causal interpretation of data, theory and are related to Campbell’s criteria for evaluating research design and the strength of research as a basis of causal inference - a universally acknowledged benchmark.

Explicitly stating the operationalization of the study variables is fundamental to interpretation. Meta-analysis cannot effectively include attributes of methodology in a moderator analysis when those methodological attributes are inconsistently reported. For example, using lineups in which the fillers (or the target substitute) are stated by the investigators to be similar to the target face begs the question how similar are they?. Subjective judgments of facial similarity have large error components. This is why there are scaling techniques using group averages rather than experimenters' impressions. Yet single person, non-quantified judgments of similarity predominate. Similarly, many ways exist to operationalize, for example, unbiased instructions, and measure, for example, lineup size and bias. Specifics are needed for interpretation.

Marketing

Claims of the superiority of sequential lineups were not made in the 1998 'White Paper' (Wells *et al.*, 1998), and were defeated in the recommendations of the Technical Working Group on Eyewitness Evidence (1999), although a strong effort was made to include the claim in the body of the document. Lindsay (1999, p. 225) made it very clear that marketing applications of ones work is an important activity, and clearly declares marketing of the sequential lineup underway. It gives us real pause to see entrepreneurial researchers mounting marketing campaigns around their research products. But, we heartily agree with the final sentence of the paper:

The application of applied research is a much more complicated process than merely conducting research that appears to address an applied issue.

The marketing process involves convincing others that one's product possesses superior attributes, which puts a strain on balanced representation of the literature. Here we present some examples.

- (1) Steblay *et al.* (2001) reported sequential lineups as superior to simultaneous lineups for overall correct decisions (56% versus 48%, $p < .0001$; $r = .09$). However, sequential lineups are superior only when the perpetrator is absent: sequential lineups are superior for minimizing false identifications of designated innocent suspects (9% versus 27%, $p < .0001$, $r = .23$) and increasing correct lineup rejections (72% versus 49%, $p < .0001$, $r = .25$). When the perpetrator is present, simultaneous lineups are superior to sequential lineups for securing correct identifications (50% versus 35%, respectively, $p < .0001$, $r = .14$) and reducing false lineup rejections (26% versus 46%, respectively, $p < .0001$, $r = .21$). Steblay *et al.* (2001, p. 468) claimed that the overall pattern of lineup accuracy supports the sequential superiority hypothesis. But this is misleading: sequential lineups are better than simultaneous lineups in some respects and worse in others. McQuiston-Surrett, *et al.* (2006) shows that the sequential advantage where the suspect is not the offender is completely offset by the simultaneous advantage where the suspect is the offender (see Malpass (2006) for a discussion of the relative importance of these outcomes).
- (2) Table 1 of Steblay *et al.* (2001) reports false identifications from two research designs: first, counting only designated suspect identifications as false IDs, and second, counting any identification in an offender-absent lineup as a false identification. Sequential lineups are reported as leading to 9% false IDs of

designated suspects whereas simultaneous lineups lead to 27% false designated suspect IDs.² The identify any foil sequential lineups lead to 28% false identifications whereas simultaneous lineups lead to 51% false identifications. These results are out of scale: sequential presentation leads to 28% false IDs (of any lineup member) versus 9% false IDs (of a designated suspect) whereas simultaneous presentation leads to 51% IDs (of any lineup member) versus 27% false IDs (of a designated suspect). A simplifying assumption that would bring these data sets into a more readily interpretable relationship is to divide the numbers for the identify any lineup member strategy by the number of lineup members, resulting in an identification number equivalent to a research strategy in which each member of the lineup were randomly designated as the suspect in the research design (Penrod, 2003).

Assuming that no individual stands out in resemblance to the actual offender, this is a reasonable strategy. Nearly all lineups in this set of studies had six members. Dividing the total number of lineup choices by 6 gives a close approximation to a randomly designated suspect result: Sequential lineups then lead to 4.67% false identifications whereas simultaneous lineups lead to 8.5% false identifications. This raises an interesting set of interpretive questions. Why is the sequential false identification rate for any foil (4.7%) approximately 1/2 of that for the designated suspect strategy (9%), and why is the simultaneous false ID rate for any foil (8.5%) approximately equal to the rate for sequential designated suspect (9%) and less than 1/3 of the simultaneous rate for designated suspect (27%)? This leads to questions about the structure of designated suspect lineups, and the worst case scenario strategy. Clark *et al.* (2008) note that the literature comparing simultaneous and sequential lineups is distinct from the general eyewitness literature in ways that suggest some differentiating factors.

Identifying a simplifying assumption for the interpretation of the false identification data for the identify any lineup member clarifies the nature of these findings for non-technical readers (consumers of the research for policy purposes) and doesn't leave them with the erroneous impression that simultaneous lineups lead to alarmingly high (51%) false identification rates.

- (3) Nearly half of the 30 tests in the Steblay *et al.* (2001) meta-analysis were unpublished articles and/or conference presentations. Including such work as part of the policy base literature (and in the meta-analysis) is problematic because unpublished articles and conference presentations have not been subjected to the more rigorous review and evaluation standards of publication, and are not archived and thus accessible for public scrutiny now and in the future. Deffenbacher, Bornstein, Penrod, and McGorty (2004) put it directly:

No unpublished studies were included, because the legal standards for proffered scientific testimony established by the US Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993, p. 692) 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.

² McQuiston et al. (2006) discuss the special status of designated suspects in target-absent identification studies.

Conclusion

If we are to give psychology away (Miller, 1969) it is important that what we give away is accurate. The reasons for this are obvious, especially when law enforcement organizations will be forced (sometimes by legislation) to comply with the dictates of psychological gifts. Criminal justice does not have its own research and development wing (Geller, 1997) and in most cases is not equipped to do its own evaluation of procedures. So it is important that we get it right.

Attempts to find alternative technologies are laudable, and the work on the sequential lineup is pioneering. However, research has not shown it to be better than what it intends to replace. Studies of lineup administration techniques in field settings have been done and more are underway. However, they face difficult obstacles for useful interpretation (Malpass & Ross, 2008; Ross & Malpass, 2008). Finding superior eyewitness identification techniques for law enforcement agencies is a work in progress.

References

- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied*, *14*, 118-128.
- Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior*, *29*(5), 575-604.
- Clark, S. E., & Davey, S. L. (2005). The target-to-foils shift in simultaneous and sequential lineups. *Law and Human Behavior*, *29*(2), 151-172.
- Clark, S. E., Howell, R. T., & Davey, S. L. (2008). Regularities in eyewitness identification. *Law and Human Behavior*, *32*, 187-218.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago, Illinois: Rand McNally.
- Daubert v. Merrell Dow Pharmaceuticals (1993). 509 US 579.
- 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*(6), 687-706.
- Ebbesen E. B., & Flowe H. (2002). *Simultaneous v. sequential lineups: What do we really know?* Retrieved from <http://www.psy.ucsd.edu/Neebbesen/SimSeq.htm> on 15 October 2008.
- Geller, W. A. (1997). Suppose we were really serious about police departments becoming learning organizations? *National Institute of Justice Journal*, *23*(4), 2-8.
- Gronlund, S. D. (2005). Sequential lineup advantage: Contributions of distinctiveness and recollection. *Applied Cognitive Psychology*, *19*, 23-37.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: SAGE Publications.
- 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. L. (1999). Applying applied research: Selling the sequential line-up. *Applied Cognitive Psychology*, *13*, 219-225.
- 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., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, *70*, 556-564.
- Malpass, R. S. (2006a). A policy evaluation of simultaneous and sequential lineups. *Psychology, Public Policy and Law*, *12*(4), 394-418.
- Malpass, R. S. (2006b). Notes on the illinois pilot program on sequential double-blind identification procedures. *Public Interest Law Reporter*, *11*(2), 5-47 (partial pages).

- 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 Malpass, R. S. (Eds.), *Simultaneous or sequential lineups? New data, new perspectives, new looks at old questions, and a study space based policy analysis*. Symposium presented at the meetings of the American Psychology & Law Society, Edinburgh.
- Malpass R. S., & Ross, S. J. (2008, June) *Implications of existing field and archival studies for field study design*. Presented at the meeting of the eyewitness identification study group. National Institute of Justice, Washington, D.C.
- Malpass R. S., Tredoux C. G., & McQuiston-Surrett D. E. (2007). Lineup construction and lineup fairness. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology (Vol. II): Memory for people*. Lawrence Erlbaum & Associates.
- Malpass, R. S., Tredoux, C. G., Schreiber Compo, N., McQuiston-Surrett, D., MacLin, O. H., Zimmerman, L. A., & Topp, L. D. (2008). Study space analysis for policy development. *Applied Cognitive Psychology*, 22(6), 789-801.
- McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006). Sequential versus simultaneous lineups: A review of methods, data, and theory. *Psychology, Public Policy and Law*, 12(2), 137-169.
- Meissner, C. A., Tredoux, C. G., Parker, J. E., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition*, 33, 783-792.
- Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist*, 24, 1063-1075.
- Penrod S. (2003). *Eyewitness identification evidence: How well are witnesses and police performing?* American Bar Association Criminal Justice Section. Retrieved 7/4/08 from <http://www.abanet.org/crimjust/spring2003/eyewitness.html>
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Revised edition). Newbury Park, CA: Sage Publications, Inc.
- Ross, S. J., & Malpass, R. S. (2008). Law and Human Behavior. *Moving forward: Response to 'Studying eyewitness investigations in the field'*, 32(1), 16-21.
- Ross S. J., & Malpass R. S. (2008, March). *Improved methodological standards for field evaluations of eyewitness identification procedures*. Paper presented at the meetings of the American Psychology-Law Society, Jacksonville, FL.
- Stebly N. (2007, March). *2001 + 6: An updated meta-analysis of eyewitness lineup performance under sequential versus simultaneous formats*. Paper presented at the 'Off the witness stand: Using psychology in the practice of justice' meeting, New York, NY.
- Stebly, 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.
- Technical working group on eyewitness evidence (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, D.C.: National Institute of Justice (i-x, 1-44). URL: <http://www.ncjrs.org/pdffiles1/nij/178240.pdf>
- 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., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photo spreads. *Law and Human Behavior*, 22(6), 603-647.
- Williams, G., & Hammelman, H. A. (1963). Identification parades. *Criminal Law Review*, 479-482, 545-555.
- Wogalter, M. S., Malpass, R. S., & McQuiston, D. E. (2004). A national survey of police on preparation and conduct of identification lineups. *Psychology, Crime and Law*, 10(1), 69-82.