Simultaneous and Sequential Lineups[1]

Roy S. Malpass
University of Texas at El Paso

The current state of our knowledge.

Studies published in excellent scientific journals over the last 2-3 years, using Receiver Operating Characteristic (ROC) analysis, a long established methodology, widely adopted across a number of scientific disciplines, indicate that if either simultaneous and sequential lineups are reliably different in proportion of accurate identifications it is likely that simultaneous lineups will be generally superior. Sequential lineups have long been known to have been based on a shaky scientific foundation, with few studies, inadequate reporting, flawed methodology and poor coverage of the research domain. Sequential lineups have been shown to provide an identification decision criterion shift as part of the presentation sequence that is difficult to control in application. The most recent studies demonstrate these outcomes plus a heightened confidence in the accuracy of false identifications for sequentially presented photospreads. Sequential lineups have not been proved generally superior to simultaneous lineups.

History.

I think it is common that researchers pursue our own lines of interest and, apart from editorial duties, often do not deeply examine lines of research undertaken by colleagues that do not intersect with their own. During the 1980's and early to mid 1990's the

---

development of sequential was something that I was content to leave to my colleagues who had established the line of research. During the deliberations of the National Institute of Justice Technical Working Group on Eyewitness Evidence [16] in 1998-1999, advocates of the superiority of sequential lineups pushed for a recommendation from the Working Group enshrining sequential lineups as superior to the more traditional [18] simultaneous form. The audience for the argument consisted of lawyers and police officers, who had no independent means of assessing the merits on their own. In the end both lineup forms were included in the published document, with no preference expressed. I went home and began a much more detailed and specific reading of the foundational studies, as did the rest of my lab.

The first meta-analysis from the Advocate community [14] was under way. The authors were helpful in pointing us to papers that we might otherwise have missed, especially providing copies or notes on unpublished studies from the lab at Queens University in Ontario. We began with the studies in the 2001 meta-analysis, and any others we could find.

In our review process we first had to grapple with some questions about how to think about reviewing an entire literature on a topic for the purpose of informing decisions about both what the literature says and about its relevance for policy. We thought of these as adequacy criteria [10, 11].

The first adequacy criterion focuses on the studies, their publication status and their distribution over laboratories. Based on the citation list of the 2011 meta-analysis [15], and combining studies in the two types of analysis the authors employed, the state of the literature as of the end of 2001 can be summarized as follows. There were 26

studies over a 17 year period, 12 (46%) unpublished. Eighteen (69%) were from the Advocate community, and sixteen of these (62% of the total studies) were from one laboratory at Queens University. Studies from the Queens lab failed to counterbalance suspect position and instead placed him in position 5. Recent research [4, 12] has shown that when the comparison dimension is correct and false identifications, placing the suspect in position 5 results in a sequential advantage in contrast to other lineup positions. The literature prior to 2001 is quite small, with nearly half of it unpublished. These are serious weaknesses in a literature offered as a policy base.

The second adequacy criterion focuses on including and reporting important aspects of methodology.

- The manner of lineup construction is an important variable in identification studies. The lineup construction method was unreported in 44% of experiments.

- It is also important to know the degree of bias towards or away from the suspect and the adequacy of the fillers as a check on lineup construction. These manipulation checks were unreported in 85% of experiments.

- It is important to calibrate identification rates with stimulus event duration, and other study characteristics to avoid restriction of range, ceiling or floor effects, etc.. Lineup-event calibration was unreported in 95% of experiments.

- Culprit-suspect similarity is also important to know in identification studies. This was unreported in 39% of experiments.

- The presence or absence of blind lineup administration is important. It is absent or unreported in 33% of experiments.

- The stopping rule is important in sequential presentations, however the stopping rule was unreported in 54% of experiments.

- Back loading (adding photos to those contained in the simultaneous lineup at the end of the sequential lineup) is an important part of the sequential lineup package. It is unreported in 44% of experiments.

Given that in large proportions of this literature we don't know how the lineups were formed, the degree of suspect-culprit similarity, the degree to which the fillers fulfill their function as alternatives to the suspect, and the degree of its bias towards the suspect, it is difficult to accurately interpret the findings.

There has been discussion about the place for narrative reviews of individual studies in the face of meta-analysis. The case of simultaneous v. sequential lineups and the first meta-analysis is an argument in favor of a more analytical treatment of individual studies forming the corpus of the literature on this topic, specifically with regard to the underlying adequacy of the literature, something meta-analysis cannot easily engage.

These reporting failures and the failure to detect and respond to the failure of the Queens laboratory to counterbalance suspect position in the sequential lineups represents a failure of the Editorial and Peer Review processes to preserve appropriate standards of research. Important questions were not asked.

The useful literature after considering these difficulties is even smaller. The studies and uncertainties embedded in the first 17 years of research on this topic form the foundational material of this literature and both their data and their limitations persist in subsequent analyses, although the latter are rarely acknowledged. Meta-analysis is a crude tool for this purpose.

There is no such thing as ***The*** Sequential Lineup. It is part of a package of factors that was present from the first study: sequential presentation of faces, double blind administration, a yes/no decision for each lineup member on first viewing, and the witness is not informed of the total number of faces to be viewed. [7, 8]. In application these additional components are widely modified or omitted. Sequential presentation is confounded with other components of the package, and there is no empirical evidence that sequential presentation is the active ingredient. Would asking a specific identification question of each member of a simultaneous photospread have the same result? We do not know. While there have been individual studies in a number of laboratories [17, 13,10], a program of research to identify and evaluate a range of procedural modifications to lineup effectiveness is lacking.

The period from 2001 to the present contained a series of meta-analyses which, taken together could not definitively resolve the question of whether or not sequential lineups were then to be considered superior to simultaneous lineups. The first was the 2001 meta-analysis authored by members of the advocate community. The analysis was controversial partly because of the relatively non-analytic nature of the underlying literature, which no meta-analysis can combat. A subsequent meta-analysis appeared in 2006 [11]. It was coupled with discussion of many of the critical points noted above and detailed examination of the individual studies contained in the 2001 meta-analysis. This meta analysis, among other things, showed that when the flawed studies of the Queens laboratory were dropped from the analysis sequential and simultaneous lineups were not statistically different and almost exactly equal.

The period from 2001 to present is configured very differently from the earlier. During these years, among the studies identified by the 2011 advocates meta-analysis there were 23 studies, 9 (39%) advocate papers and 14 (61%) from independent laboratories. The 2011 meta-analysis found sequential superiority, but has been severely criticized for devising exclusion criteria resulting in the exclusion of one large scale study [4] which found few differences between simultaneous and sequential lineups, and those that were found balanced in their superiority contingent on the position of the suspect in the sequential lineups. This 2009 study was rigorous in measuring, controlling and reporting the study attributes discussed above, vastly improving its interpretability. A 2012 paper [2] analyzed 51 tests of the superiority of one lineup format over the other and failed to find significant differences, except offsetting effects in that sequential lineups produced both fewer correct identifications and fewer false identifications, reflecting a more conservative decision criterion for the sequential procedure.

The bulk of research on simultaneous and sequential lineups has been what program evaluation researchers would call "summative" studies, which are concerned, in this case, with which "program" is superior with respect to particular criteria. There has been remarkably little analytic study that evaluators might call "formative" evaluation - that is studies carried out with the purpose of finding identification procedures that optimize outcomes under widely varying conditions of witnessing and identification procedure administration.

A major activity during the second period in the life of sequential lineups has been the search for more useful dimensions of evaluation and comparison. The obvious and

historical ones are correct and false identifications. But these call attention to the question of whether our work produces techniques that favor correct IDs (the prosecution) or reduce false IDs (the defense). While it is essential to know these things, they are superficial. It is perhaps more important to develop indicators that are related more closely to theory about the psychological process involved and that incorporate policy considerations. There are at least 6 additional numerical indices [2] and another set of evaluative dimensions based on the utilities of expected outcomes [2, 9]. It would be difficult to choose among these because they have different sensitivities to different aspects of the identification process and different implications for how to improve the system. However there is another approach that is both more comprehensive than this set of indicators and brings the field of eyewitness identification into more intimate contact with decades of important research in memory and perception, and response decision making processes which embody personal values and likelihood estimates.

The most significant development in research on eyewitness identification procedures, from showups to multi-person lineups of any kind, is the application of Receiver Operating Characteristic (ROC) analysis [12, 6], which we have heard about more extensively earlier today. Suffice it to say that the application of ROC analysis to the eyewitness identification problem is of truly seismic proportions for our field. Already researchers are re-analyzing earlier findings [5] and contributing new research based in ROC analysis [3]. The results are different from earlier findings based on raw indices such as correct IDs, false IDs and their ratios.

So what is the implication of ROC analysis for the question of sequential and simultaneous lineups? Early returns indicate that in the laboratory sequential lineups are sometimes as effective as simultaneous lineups but not better and most frequently worse. We have known ever since researchers outside of the advocate group began closer examination of the literature that sequential lineups were probably not better than simultaneous lineups when considering both correct and false identifications, and allowing for base rate and other variations [2, 9]). I think we now know something more, and considerably more deeply. This work has essentially set aside much of the earlier work. Is the effect of biased v. unbiased instruction to move the criterion on a single ROC or does it reveal different ROCs? And, does it change the recommendation to use unbiased instructions. Only reanalysis, or new studies will reveal the answers.

The insight that sequential and simultaneous lineups are on different ROCs and the possibility that the same situation may help to understand things like the own-group bias effect is one of the greatest aesthetic experiences I have had in my decades as a scientist. I suppose I should have figured that out while working on my cross-race recognition paper in 1968-69 when I used a detection theory approach. But I didn't get to the eyewitness paradigm for another 10 years. My preparation in ROC analysis was not up to the insight. I am happy almost to tears to see it arrive.

In my view, the literature claims to support sequential superiority is based on inadequate science, flawed by methodological errors, the small number of studies actually published, confusion about the policy relevant dimensions of comparison and the domination of the research corpus by advocates. I have never been convinced that sequential lineups were superior, as a matter of public policy. Now that ROC analysis

has discovered the various phenomena of eyewitness identification and given us elegant and powerful analytic schemes that are broadly useful and ripe with interesting implications, I think that much of what we thought we knew will have to be revisited and reanalyzed. The various meta-analyses can now be seen as largely uninformative because the dimensions of comparison used previously are inadequate. I am convinced by the recent publications using ROC analysis that sequential lineups are not superior. And since their associated demands and vulnerabilities are real and difficult to control in application, The jurisdictions that have required their use should delete the preference from their policy and procedure documents. A concentration on high quality lineups and more carefully trained and supervised administration of identification procedures would be a far better use of time and resources.

Regarding continuing advocacy of and proselytizing for sequential lineups as a superior form of identification, that ought to stop, because the science does not support it. I agree with the suggestion of a recent article [3] that jurisdictions which have mandated sequential lineups should rethink that decision.

So, have these last 18 years been a waste of time, energy and money? I don't think so. The close attention to the comparison has brought more sophisticated attention to theory, measurement and context, and we have discovered ROC analysis (or, has it discovered us?), which has all the time been alive and well in cognate fields that we really should have paid more attention to long before this. We should embrace it and move forward.

Citations

1. Carlson, C. A., Gronlund, S. D., & Clark, S. E. (2008). Lineup composition, suspect position, and the sequential lineup advantage. *Journal of Experimental Psychology: Applied*, *14*(2), 118-128. doi:10.1037/1076-898X.14.2.118

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

3. Dobolyi, D. G., & Dodson, C. S. (in press, 2013, November 4). Eyewitness Confidence in Simultaneous and Sequential Lineups- A Criterion Shift Account for Sequential Mistaken Identification Overconfidence. *Journal of Experimental Psychology- Applied.*

4. Gronlund, S., D., Carlson, C. A., Dailey, S. B. & Goodsell (2009). Robustness of the Sequential Lineup Advantage. *Journal of Experimental Psychology-Applied, 15(2)* 140-152

5. Gronlund, S.D., Carlson, C.A., Neuschatz, J. S., Goodsell, C.A., Wetmore S.A., Wooten, A. & Graham, M. (2012). Showups versus lineups- An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition.*

6. Gronlund, S.D., Wixted, J. T. & Mickes, L. (in press). *Current Directions in Psychological Science.*

7. Lindsay, R. C. L., & Bellinger, K. (1999). Alternatives to the sequential lineup: The importance of controlling the pictures. *Journal of Applied Psychology, 84*, 315–321. doi:10.1037/0021-9010.84.3.315

8. Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*, 556–564. doi:10.1037/0021-9010.70.3.556

9. Malpass, R. S. (2006). A Policy Evaluation of Simultaneous and Sequential Lineups. *Psychology, Public Policy and Law, 12(4)*, 394-418

10. Malpass, R. S., Tredoux, C. G., Schreiber Compo, N., McQuiston-Surrett, D. E., MacLin, O. H., Zimmerman, L. A. & Topp, L. D. (2008). Study Space Analysis for Policy Development. *Applied Cognitive Psychology, 22(6)* 789-801

11. McQuiston-Surrett, D. E., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. Simultaneous Lineups- A Review of Methods, Data, and Theory. *Psychology, Public Policy and Law. 12(2)*, 137-169.

12. Mickes, L., Flowe, H., & Wixted, J. T. (2012). Receiver Operating Characteristic Analysis of Eyewitness Memory- Comparing the Diagnostic Accuracy of Simultaneous Versus Sequential Lineups. *Journal of Experimental Psychology- Applied, 18(4),* 361–376

13. Pozzulo, J. D., Corey, S. C., Kiraly, M., Girardi, A., Lawandi, A., & Aston, C. (2008). Can a lineup procedure designed for child witnesses work for adults: Comparing simultaneous, sequential, and elimination lineup procedures. *Journal of Applied Social Psychology, 38,* 2195–2209. doi:10.1111/j.1559-1816.2008.00387.x

14. Steblay, N. K., 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. doi:10.1023/A:1012888715007

15. Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. Psychology, Public Policy, and Law, 17, 99–139. doi:10. 1037/a0021650

16. 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)

17. Weber, N., & Perfect, T. J. (2012). Improving eyewitness identification accuracy by screening out those who say they don't know. *Law and Human Behavior, 36*, 28–36.

18. 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.