

A LINEUP EVALUATION “DO-IT-YOURSELF KIT” for Attorneys and Law Enforcement.

Roy S. Malpass
Eyewitness Identification Research Laboratory
University of Texas at El Paso

Procedures for constructing a fair lineup are worked out and available elsewhere. For a summary, go to our discussions of *General Principles*¹ and *Evaluating Lineup Fairness*². A detailed presentation of procedures for implementing a quantitative evaluation of lineup fairness is normally left out of these discussions. The present document is an attempt to provide an entry-level guide to lineup evaluation, and to provide some of the basic tools researchers use in the process. An introductory guide to these can be found in Malpass & Lindsay’s (1999) introduction to the Special Issue of Applied Cognitive Psychology devoted to “Measuring Lineup Fairness”³.

Introduction to Evaluating Eyewitness Lineups - Yours and Theirs.

A. The purpose of the Lineup Evaluation Kit is to determine two things:

1. Does the suspect stand out in the lineup?
2. Are the fillers adequate alternatives to the suspect?

Both of these questions are answered with the same basic procedure.

B. The procedure involves four steps.

1. Obtain the description of the suspect given by the witness(s).
2. Type the verbal description on a single page, with no other information on the page.
3. Identify a set of people of approximately the same demographic characteristics as the witness. These will be your “mock witnesses”.
4. Individually, and in a setting where you will not be interrupted,
 - a. tell each mock witness in turn that they will be shown a description of an offender that was given by a witness to a crime.

If there is no useful description given by an eyewitness, do this: for about 10 seconds, show a photo of the suspect to 3 or 4 persons of the general demographic characteristics as the witness (same sex, same race). Take away the photo, and ask them to describe the person in the photo. Compile a composite description by retaining those elements of the description given by 2 or more of them. Use this as the description from which the lineup should be constructed and evaluated.
 - b. Show the description to each mock witness, and give him/her ample time to study the description.
 - c. Take the description page from the mock witness, and
 - d. Show the mock witness the lineup and ask her/him to tell you, based on the description they were given, which member of the lineup is the offender.
 - e. Record the location indicated by the mock witness. In the case of a sequentially presented lineup, identify and record the specific photo indicated by the mock witness
 - f. Repeat this process for at least 20 mock witnesses.

¹ <http://eyewitness.utep.edu/consult04.html>

² <http://eyewitness.utep.edu/consult05.html>

³ <http://eyewitness.utep.edu/M&L.pdf>

The Tests.

Your record of the 20+ mock witness judgments will serve as a basis for testing two aspects of lineup fairness:

1. Lineup **bias** towards or away from the suspect/defendant in this investigation,
2. Lineup **size**: The degree to which the lineup fillers are adequate alternatives to the suspect.

A. Bias.

The spreadsheet "Lineup Bias"⁴ allows you to enter summary data from your mock witness assessment and calculates certain important figures. First, enter your summary data as in the example in Figure 1.

TEST FOR LINEUP BIAS	
Instructions.	
1. Enter the total number of mock-witness choices here:	29
2. Enter the number of mock witnesses who chose the suspect here:	14
3. Enter the total number of persons in the lineup here:	6
Note: Fields in which data can be entered are shown in boxes. Other fields are calculated fields.	

Figure 1. Test for Lineup Bias: Data Entry

1. Enter the number of mock witnesses who made a choice in the lineup. In Figure 1 there are 29 such people.
2. Enter the number of mock witnesses identifying the suspect/defendant. In Figure 1 there are 14 such people.
3. Enter the number of persons/faces in the lineup. In Figure 1 there are 6 photos in the lineup.

The spreadsheet will allow you to enter numbers in these three fields only. All others are locked, and will only display calculated values derived from these three data entries. The spreadsheet calculates a series of statistics, but the two that are of interest are these:

- The proportion expected to identify the suspect if chance alone determines the choices, and
- The proportion of mock witnesses actually identifying the suspect

Testing the reliability of the discrepancy between the observed and expected values.

The magnitude of the discrepancy is associated with a specific probability of occurrence through chance alone. In the present example, the proportion of identifications of the suspect is .483, or 48.3%. The proportion expected if only chance determines the mock witness's choice is 1/6, .167, or 16.7%, and the difference between them is .316, or 31.6%. In general, we can calculate the probability that any specific difference between observed and expected choice rates will occur by chance alone. In practice, however, we are interested in whether the probability of a particular difference is below conventional levels of probability. If the critical ratio (see below) exceeds 1.96, the difference is likely to occur only 5% of the time if only chance factors are operating.

If the ratio exceeds 2.58, the difference is likely to occur only 1% of the time if only chance factors are at work. It is too much for the present document to explain the statistical basis of these critical values, but for those who are interested nearly any introductory book on basic statistics will provide that information.

In the present example, the critical ratio is 3.406, far beyond 2.58, and indicating that a difference of this magnitude is extremely unlikely to occur by the operation of chance alone. It is very likely that some biasing factor has directed the mock witness's attention to the suspect, producing a biased lineup. For an example, see our webpage on the importance of photo choice.

⁴ <http://eyewitness.utep.edu/bias-calc.xls>

Proportion identifying the suspect (p)	.483
q = 1 - p	.517
n	29
Standard Error of p	.093
Chance expectation:	.167
Critical Ratio for diff from Chance:	3.406
(.05 requires 1.96)	
(.01 requires 2.58)	
If the proportion of mock witnesses identifying the suspect reveals bias, beyond a chance probability of .05, the critical ratio will exceed 1.96.	
If the proportion of mock witnesses identifying the suspect reveals bias, beyond a chance probability of .01, the critical ratio will exceed 2.58.	

Figure 2. Test for Lineup Bias: Calculation of Bias and its reliability

This is not “rocket science”! These techniques are simple, standard, and found in nearly any elementary text in statistics and research design.

Another way to use the mock witness data is to calculate “confidence intervals” around the estimate of the proportion identifying the suspect / defendant. In particular, it is important to note whether the identification rate expected through the action of chance alone (.167, or 16.7%) is excluded from the confidence interval around the identification rate estimate. In the example in Figure 2, the estimate of the actual identification rate is associated with specific amounts of uncertainty.

Confidence intervals			
5% confidence interval equals	0.301	to	0.665
1% confidence interval equals	0.243	to	0.722
If the proportion of mock witnesses expected to identify the suspect by chance is outside of the 5% confidence interval, we can be 95% certain that the lineup is biased.			
If the proportion of mock witnesses expected to identify the suspect by chance is outside of the 1% confidence interval, we can be 99% certain that the lineup is biased.			

Figure 3. Identification confidence intervals

Calculating confidence intervals allows us to establish boundaries around our findings (like the margins of error in political polls) and calculate the probabilities of reaching beyond these boundaries. Given the error inherent in the estimated identification rate of .483 (48.3%) because of the small number of mock witnesses on which the estimate was based, we can be 95% certain that the true mock witness identification rate for this lineup is between .301 at the low end, and .665 at the high end. We can be 99% confident that the true mock witness identification rate for this lineup is between .243 at the low end and .722 at the high

end. The importance of these figures is that they exclude the chance-only value of .167: This lineup is at least biased to the extent of .243 (24.3%) identifications, above the theoretically unbiased figure by 45%. At most, this lineup is biased to the extent of .772 (77.2%) identifications, above the theoretically unbiased figure by 362%, with the most likely figure being the observed identification rate of .483, placing the most likely bias at 189% above the theoretical chance level. It is appropriate to conclude, on the basis of the evidence collected in the evaluation of this lineup, that it is biased against the suspect / defendant. We cannot exactly determine the source of bias, but it seems clear that it is present.

There are some conditions under which you may wish to repeat this evaluation procedure multiple times.

- If there were multiple witnesses, who gave substantially different descriptions,
- If there were witnesses of different racial groups.

B. Size.

The spreadsheet “Lineup Size”⁵ allows you to enter summary data from your mock witness assessment and calculates certain important quantities. First, enter your summary data as in the example shown in Figure 4.

TEST FOR LINEUP SIZE	
using Tredoux' E	
Enter the total number (N) of mock witnesses here:	29
Enter the total number (K) of persons in the lineup here:	6
Note: Fields in which data can be entered are shown in boxes.	
Other fields are calculated fields.	

Figure 4. Test for Lineup Size: First Data Entry

1. Enter the number of mock witnesses who were shown the lineup. In the example there are 29 such people.
2. Enter the number of persons/faces in the lineup. In the example below there are 6 photos in the lineup.

Then enter the number of mock witnesses who identified each member of the lineup (the spreadsheet allows for up to 10 lineup members – see Figure 5).

The spreadsheet calculates a series of statistics, but the one that is of interest is known as Tredoux' E. Tredoux' E estimates the number of persons in the lineup who are realistic choices given the verbal description of the offender, whether this is obtained from witness descriptions of the offender or descriptions you produce based on the suspect for the purposes of carrying out this evaluation.

Figure 5 shows that there are just over 3 members of the lineup who are realistic choices given the offender/suspect description – including the suspect. But before we get into the calculations, let's look at the distribution of choices in the table. There were six members of the lineup, and 29 mock witnesses. That means that by chance alone we would expect $29/6 = 4.83$ IDs for each member of a lineup in which each filler is an ideal alternative to the suspect (according to the verbal description). Looking at the table, we see that 3 members of the lineup are chosen at or above this frequency, and one was chosen considerably more often. This is what Tredoux' E measures with great precision. But from a practical point of view, it is pretty easy to see which lineup members are not adequate choice alternatives, and which ones are super-attractive to mock witnesses.

If you are constructing a lineup, and using mock witness procedures to check your work, the basic table actually tells you what you need to know: that certain fillers fail to draw choices from mock witnesses, or

⁵ <http://eyewitness.utep.edu/size-calc.xls>

that certain fillers draw too many (a “super” filler). There is, as you may have guessed, a dynamic relationship between fillers. Replace a super-filler with another one and others who did not attract choices before may attract them now.

First lineup member	2
Second lineup member	0
Third lineup member	2
Fourth lineup member	6
Fifth lineup member	5
Sixth lineup member	14
Seventh lineup member	0
Eighth lineup member	0
Ninth lineup member	0
Tenth lineup member	0
Sum = N, =	29

Figure 5. Test for Lineup Size: Second Data Entry

Tredoux’ E calculates that the size of this lineup is 3.174 effective members. And if that’s all Tredoux’ E did for us, it would be a complicated way of determining what is usually clear from the table of mock witness choices. But in addition, Tredoux’ E provides us with a way to calculate a confidence interval around this number. As with the bias figure (the proportion of mock witnesses identifying the suspect) we can calculate the lineup sizes that can be excluded based on the value of Tredoux’ E and the number of mock witnesses used.

Tredoux' E	3.174		
95% confidence interval equals:	2.290	to	5.165
99% confidence interval equals:	2.105	to	6.445

Figure 6. Tredoux’ E, where the number of mock witness equals 29

As with political polls, there is a margin of error. Imagine a scale from zero to 10. In the present instance it is most likely that the true size of the lineup is 3.174, and it is only 5% likely that the true size is greater than 5.165 or less than 2.29. It is even less likely (1%) that it lies above 6.445 or below 2.105. No value is more likely than 3.174. From a practical point of view, we know that the size is most likely to be 3 (that is, there are only 3 lineup members who are realistic choices for mock witnesses, based on the verbal description). And we know that it is very unlikely that the number of realistic choices is 2 or fewer, or greater than 5 (figure 6).

The margin of error depends on the number of people (mock witnesses) who make lineup choices. If we use 5 X the total number of mock witnesses but keep the same proportion of choices for each of the individual lineup members, we get the results shown in figure 7.

With the increased number of mock witnesses, we now know that the most likely number of realistic choices is still 3, but we have refined the limits. We know that it is unlikely that the true number is 2 or below, or 4 and above (according to the 5% level of confidence) or 5 and above (according to the 1% level of confidence).

So in the end, we know what looking at the table told us, but we also know pretty clearly how much confidence we can have in that knowledge.

	Tredoux' E	3.17		
	95% confidence interval equals:	2.71	to	3.83
	99% confidence interval equals:	2.59	to	4.11

Figure 7. Tredoux' E, where the number of mock witnesses = 145

Summary.

So now, what do we have? We know that we can evaluate lineups that we have put together with a simple technique that calls for 2-3 dozen people to judge the lineup. And from these judgments we can know whether we have reached a level of fairness that is defensible under close critical examination. If we have not, we know where to start to substitute better fillers, and after another iteration of the process it is very likely that we will succeed.

We also know that we can evaluate the lineups others have constructed, and we will have quite a bit of detailed knowledge of the strength or weakness of their product. This is a useful tool for training as well as for legal processes.

Bibliography.

If you wish to read further in the literature on the use of mock witnesses in measuring lineup fairness, please consult our bibliography on this topic at <http://eyewitness.utep.edu/mock.html>.