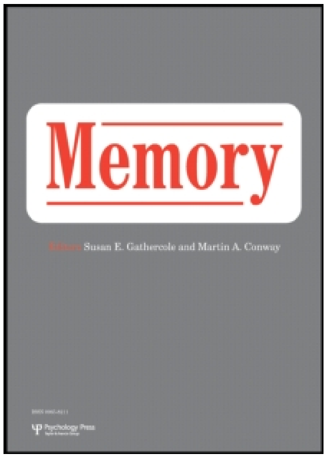

Memory
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683358

The impact of eyewitness identifications from simultaneous and sequential lineups

Daniel B. Wright [a]
[a] University of Sussex, Brighton, UK

Online Publication Date: 01 October 2007
To cite this Article: Wright, Daniel B. (2007) 'The impact of eyewitness identifications from simultaneous and sequential lineups', Memory, 15:7, 746 - 754
To link to this article: DOI: 10.1080/09658210701508401
URL: http://dx.doi.org/10.1080/09658210701508401

MEMORY, 2007, 15 (7), 746–754

Psychology Press
Taylor & Francis Group

The impact of eyewitness identifications from simultaneous and sequential lineups

Daniel B. Wright
University of Sussex, Brighton, UK

Recent guidelines in the US allow either simultaneous or sequential lineups to be used for eyewitness identification. This paper investigates how potential jurors weight the probative value of the different outcomes from both of these types of lineups. Participants ($n = 340$) were given a description of a case that included some exonerating and some incriminating evidence. There was either a simultaneous or a sequential lineup. Depending on the condition, an eyewitness chose the suspect, chose a filler, or made no identification. The participant had to judge the guilt of the suspect and decide whether to render a guilty verdict. For both simultaneous and sequential lineups an identification had a large effect, increasing the probability of a guilty verdict. There were no reliable effects detected between making no identification and identifying a filler. The effect sizes were similar for simultaneous and sequential lineups. These findings are important for judges and other legal professionals to know for trials involving lineup identifications.

In the first edition of his popular textbook, Hilgard states: "experiments on testimony have shown that many things are remembered that never happened at all or that actually happened in ways very different from those recalled" (1953, p. 270). Many people in the justice system realise this. As Janet Reno, then US Attorney General, said in the preface to *The Guidelines* for interviewing eyewitnesses and constructing lineups: "Even the most honest and objective people can make mistakes in recalling and interpreting a witnessed event; it is the nature of human memory" (Technical Working Group for Eyewitness Evidence [hereafter TWGEE], 1999, p. iii). *The Guidelines* were produced because of a growing public awareness of how the fallibility of human memory affects the criminal justice system. They are important because they describe how to interview eyewitnesses and how to conduct a lineup. Other countries have different rules and guidelines for conducting lineups. For example, in the UK (where lineups are called identification parades), the Police And Criminal Evidence (PACE) act describes how lineups should be conducted. The most recent version is available on http://police.homeoffice.gov.uk/operational-policing/powers-pace-codes/pace-codes.html. Kebbell (2000) describes some of the differences between the US and the UK documents.

Surveys of London lineups have found that about 20% of the time eyewitnesses choose one of the fillers (Valentine, Pickering, & Darling, 2003; Wright & McDaid, 1996; see Behrman & Davey, 2001, for similar figures in the USA). Although

Address correspondence to: Daniel B. Wright, Psychology Department, University of Sussex, Brighton, BN1 4EP, UK. E-mail: DanW@sussex.ac.uk

Thanks to Steve Darling and Amina Memon for valuable comments on this research, and to Roland Baddeley and Graham Hole for helping with the data collection.

http://www.psypress.com/memory DOI:10.1080/09658210701508401

people acting as fillers are not arrested,[1] eyewitnesses making mistakes like these suggests that innocent suspects are also often chosen. On the basis of survey data and laboratory studies, Penrod (2003) estimates that 10–15% of the time when a suspect is identified the person is not the culprit. The consequences of this can be severe. The majority of cases where convictions were overturned by DNA evidence involved errant eyewitness testimony (Scheck, Neufeld, & Dwyer, 2003; see also www.innocence.com). This has led many people to conclude that eyewitness errors account for the majority of false convictions (see Wells & Olson, 2003; Wright & Davies, 2007, for reviews). Similarly, in about 40% of lineups the eyewitness makes no identification. In some of these lineups the suspect will not be the culprit and therefore the eyewitness is making a correct rejection. But according to Penrod (2003), about half of the time the unidentified suspect is the culprit, and therefore this is another type of error. It is necessary to consider both false identification errors and misses when evaluating lineup procedures.

The problem is not just that eyewitnesses make errors, but that people overestimate the diagnostic value of eyewitness evidence. In one of the first studies of its type, Loftus (1974) gave participants a crime scenario where a man robbed a grocery shop and shot dead two people. There was only a small amount of incriminating evidence and only 9 of 50 participants presented with this evidence rendered a guilty verdict. For a second group of 50 participants, Loftus added the information that a store clerk saw the culprit and identified the suspect. Guilty verdicts were given by 36 of the 50 participants. A third set of participants were told that the eyewitness had very poor vision, making the identification unreliable. Still, 34 of 50 participants rendered a guilty verdict. This study shows that people weight eyewitness evidence highly and that they are often uncritical of the reliability of the testimony.

It is worth describing Loftus's (1974) study in more detail in order to introduce some of the descriptive statistics that will be used to analyse the data in the current study. The most common inferential statistic in these circumstances is χ^2. In Loftus (1974) it is 36.32, which is significant, but this tells the reader neither about the size of the effect nor where differences exist. There are several ways to describe these (Agresti, 2002). One measure of the overall effect size is the size of the observed χ^2 relative to the maximum χ^2 that could be attained. The maximum attainable χ^2 is $n \times \min(r-1,\ c-1)$, where n is the total sample size, r is the number of rows, and c is the number of columns. Here the maximum χ^2 is n because it is a 3×2 χ^2. Thus, the proportion of possible χ^2 variation accounted for is $36.32/150 = .24$. The square root of this is called Cramér's V, which is provided by many mainstream statistical packages (like SPSS). Here, V^2 is used since it is easier to interpret and is a categorical analogue to the effect size measure used here for continuous data, η^2. Both V^2 and η^2 go from 0 to 1, with 0 meaning no association and 1 meaning perfect association. The odds ratio is one of the most common effect size measures to compare pairs of conditions, although it is not without fault (Kraemer, 2004). The odds are defined by the number of occurrences of some event divided by the number of times the event did not occur. For the Loftus (1974) study, the odds of a guilt verdict are the number of guilty verdicts over the number of not guilty verdicts. The odds for the no eyewitness condition are $9/41 = 0.22$ and for first eyewitness condition $36/14 = 2.57$. The odds ratio is the ratio of these: $2.57/0.22 = 11.7$. Thus, the odds of a guilty verdict are over 11 times higher if an eyewitness makes an identification than if no eyewitness evidence is presented.

The current study extends Loftus's (1974) study in two directions. First, both simultaneous and sequential lineups are used. This is important because sequential lineups are now being used in many jurisdictions. Second, there is an emphasis on comparing situations when the eyewitness makes no identification with situations when the eyewitness identifies a filler. The remainder of this introduction describes the importance of these two directions.

One of the most important aspects of *The Guidelines* (TWGEE, 1999) was that it allowed either simultaneous or sequential lineups to be conducted (Wells et al., 2000). While simultaneous lineups are still the norm, an increasing number of jurisdictions are using sequential procedures, for example, in California, New Jersey, and New York (Perrotta, 2004; Willing, 2002). In a simultaneous lineup, an eyewitness views the suspect and several fillers (people known not to be culprit) simultaneously, is asked if the culprit is present, and if so which person is the culprit.

[1] There are exceptions. See http://www.psychology.iastate.edu/faculty/gwells/FillerCharged.htm accessed 28.01.2005.

In a sequential lineup the eyewitness views the suspect and fillers one at a time. There are two main types of sequential lineup. In the original sequential lineup for each person (i.e., the suspect and fillers) the eyewitness either identifies the person as the culprit or not. If the eyewitness makes an identification the procedure ends. If no identification is made then the next person is shown to the eyewitness. Beginning with Lindsay and Wells (1985), most research comparing simultaneous and sequential lineups finds that sequential lineups produce fewer false positives than simultaneous lineups, but often there is a decrease in correct identifications (see Steblay, Dysart, Fulero, & Lindsay, 2001, for meta-analysis). There remains much debate about why these procedures produce different results (Meissner, Tredoux, Parker, & MacLin, 2005). A modified version of the sequential lineup allows the eyewitness to view all the people and either to make multiple identifications or to make the identification at the end of seeing all the people (see Levi, 1998, for related discussion). Some jurisdictions have adopted this version, although there is less research for this approach. It is effectively the approach used in the UK with video lineups.

Our question relating to the original sequential lineup is whether people evaluate identification information from simultaneous and sequential lineups in a similar way. In particular: Is the effect of eyewitness identification on jurors' beliefs as large for sequential as for simultaneous lineups? This is important given that jurors are now being faced with evidence from sequential lineups.

The second direction this research takes is differentiating between choosing a filler and making no identification. Often this is not done by the police conducting the lineup, which is unfortunate because these situations provide different diagnostic information. Wells and Olson (2002) show that a filler identification is more indicative of the suspect being innocent than when no identification is made.

Two measures of the participants' beliefs are taken: verdict and belief in guilt. This provides two measures of the impact of the identification, and also allows us to explore the relationship between belief in guilt and verdict for the different conditions. There were several other methodological choices made. First, only a short crime description was used (see Appendix) and participants did not discuss the case with others, as happens with real juries. The interest here is how the identification affects individual judgements, and therefore it was important not to include too much additional judicial material nor to include group discussion. Clearly interactions between the impact of an identification and these additional factors are important to explore, but in this initial exploration the focus was on the simpler effects and without discussion. Second, students were used. This was because they are relatively easy to recruit and it has the benefit that the participants in the different conditions will be roughly matched for educational level. Finally, only one crime situation was used. In this sense, this is a case study (Wells & Windschitl, 1999), and other researchers are encouraged to investigate how the impact of an identification varies by different types of crime situations.

METHOD

A total of 340 undergraduate students volunteered to take part in this study during a laboratory session. The study took less than 10 minutes and none of the participants was paid. Almost all are British citizens and eligible for jury duty. They read a description of the case of *Nancy Von Roper* (real names not used), which originally appeared in Loftus and Ketcham (1991). It involved eyewitness misidentification and police corruption, leading to the false conviction of *Tom Hoyle*. Evidence was selected from the case to allow some variability in people's belief in guilt. Both exonerating and incriminating evidence was included (see Appendix). Earlier research (Wright & Hall, 2007) shows that this set of evidence should produce approximately 50% belief in guilt judgements when the suspect is identified. The seven conditions were:

1. Simultaneous lineup with suspect chosen by eyewitness.
2. Simultaneous lineup with a filler chosen by eyewitness.
3. Simultaneous lineup with no identification made.
4. Sequential lineup with suspect chosen by eyewitness.
5. Sequential lineup with the eyewitness viewing the suspect, but choosing a filler.
6. Sequential lineup with the eyewitness choosing a filler before viewing the suspect.
7. Sequential lineup with no identification made.

Participants in the simultaneous conditions were told: "They showed all six photographs at the same time." Participants in the sequential conditions (the wording here is for condition 4, the other conditions had the wording changed appropriately for the condition) were told:

> They showed the photographs one at a time. The first three photos shown were of other people and Nancy Von Roper said that none of them were the person who raped her. The fourth photograph was of Tom Hoyle. Nancy Von Roper identified Tom Hoyle's photograph as the person who raped her. The identification procedure ended; Nancy Von Roper was not shown the final two photographs.

The number of people viewed during a sequential lineup depends on if there is an identification, and if so which person is identified. If the eyewitness identifies the first person this is the only person seen. Therefore, careful consideration was made of when to say that the witness made an identification. In condition 4 (above) the witness would have seen four photographs. In condition 5 the witness viewed a total of five photographs (including the suspect), but the fifth person was a filler. In condition 6 the witness picked the second person, so only saw two photographs. In the final condition the witness viewed all six photographs. Therefore, the number of photographs seen is not the same across conditions. This is something that occurs with sequential lineups, so it was decided not to control for this.

The sheets for the different conditions were randomly ordered and handed out prior to the participants' lectures. Participants were tested in groups of about 20. Participants read the case description and then had to make two judgements. First they had to render either a guilty or a not guilty verdict, and then they had to estimate their belief that defendant was the actual culprit on a 0–100% scale. This rating may be affected by the verdict question, but because both measures were needed for each person to explore some hypotheses, both were asked. All participants were asked these questions in the same order. The interest is in comparisons among the eyewitness conditions, so it was important to have all participants treated equally other than this manipulation. After completing the study participants were debriefed and thanked.

RESULTS

Table 1 gives the percentages of guilty verdicts and the means for belief in guilt. The data for the simultaneous lineups follow the pattern expected from Loftus (1974). The participants are more than twice as likely to render a guilty verdict if the suspect has been identified than if they have not. Table 1 also reports the omnibus test statistics (χ^2 for verdicts and F for belief) and effect sizes (V^2 for verdicts and η^2 for belief). There are statistically significant differences, for both verdicts and beliefs, between people told that the eyewitness identified the suspect versus people told that he was not identified with the

TABLE 1
The percentage of guilty verdicts and the mean belief in guilt for all seven conditions

	Simultaneous				
	Suspect (1)	*Filler (2)*		*No ID (3)*	*Test statistics*
Guilty verdicts	32/53 = 60%	17/66 = 26%		12/54 = 18%	$\chi^2(2) = 26.10$ $V^2 = .14$, $p < .001$
Belief in guilt	67.53 $SE = 2.63$	54.85 $SE = 2.88$		50.00 $SE = 2.24$	$F(2, 182) = 11.14$ $\eta^2 = .11$, $p < .001$
	Sequential				
	Suspect (4)	*Filler after (5)*	*Filler before (6)*	*No ID (7)*	*Test statistics*
Guilty verdicts	21/39 = 54%	10/34 = 29%	5/35 = 14%	11/47 = 23%	$\chi^2(3) = 15.55$ $V^2 = .10$, $p = .001$
Belief in guilt	63.38	57.53	50.83	53.63	$F(3, 150) = 2.65$
Belief in guilt	$SE = 3.08$	$SE = 2.75$	$SE = 3.95$	$SE = 3.32$	$\eta^2 = .05$, $p = .05$

The numbers in parentheses refer to the condition numbers listed in the text and used in Table 2.

TABLE 2
The effect sizes and their associated probabilities for nine pairwise comparisons

	Conditions		*Verdicts odds ratio (p)*	*Belief* η^2 *(p)*
Within simultaneous lineups	1	2	4.39 (<.001)	.08 (<.001)
	1	3	6.86 (<.001)	.18 (<.001)
	2	3	1.56 (.29)	.01 (.19)
Within sequential lineups	4	5	2.80 (.04)	.03 (.17)
	4	6	7.00 (<.001)	.08 (.01)
	4	7	3.82 (.004)	.05 (.04)
	5	6	2.50 (.13)	.03 (.17)
	5	7	1.36 (.54)	.01 (.39)
	6	7	0.55 (.30)	.00 (.59)

The odds ratios are for the verdicts; the η^2 are for the belief measures.

simultaneous lineups. For the sequential lineups, the differences were in the same direction. The effect was statistically significant for verdicts, but only marginally significant for belief.

Pairwise comparisons were done for all pairs within lineup type (see Table 2). For both verdicts and beliefs, there were three tests for the simultaneous lineups and six tests for the sequential lineups. Thus, there were nine tests for the verdicts and nine for the beliefs. Depending on several factors, including what counts as a ''family'' of tests and the relative disadvantages of increasing the chance of Type 1 and Type 2 errors, there are different ways to adjust for multiple comparisons. If the consequences of making a Type 1 error are large relative to making a Type 2 error, then the critical p value for any comparison should be lowered to, for example, .01.[2] If consequences of making a Type 2 error are more detrimental then the critical p value should be raised to, for example, .10. Table 2 reports pairwise comparisons (the odds ratio for verdicts and η^2 for beliefs, with their associated p values) without adjustment. Given the questionable value of p values (Cohen, 1994; Wilkinson & the Task Force on Statistical Inference, 1999; Wright, 2003), this seems the pragmatic approach. The pattern of results is more important than the results for any single comparison.

For the verdicts, the eyewitness making an identification had a statistically significant impact compared with the other conditions for both types of lineup. For the simultaneous lineups there was also convincing evidence that belief in guilt was affected by the identification. The evidence of this is less strong for the sequential lineups, but the effects are all consistent. Making no identification did not differ from choosing a filler for either type of lineup for both verdicts and belief.

Thus far the analyses have been done separately for simultaneous and sequential lineups. This is because of the different types of filler identifications and number of photographs seen in each. Overall, there were 60% guilty verdicts in simultaneous lineups when the suspect was identified and 22% when the suspect was not identified. The corresponding percentages for sequential lineups are 54% and 22%. To test differences a logistic regression was run to predict verdict. The interaction between lineup type and identification was nonsignificant, $\chi^2(1) = 0.31$, $p = .58$, and was removed. The main effect of lineup type was nonsignificant, $\chi^2(1) = 0.27$, $p = .60$, and the main effect of identification was significant, $\chi^2(1) = 38.87$, $p < .0001$, in the predicted direction. The odds of a guilty verdict are, overall, 4.77 times higher if the eyewitness makes an identification (95% CI from 2.86 to 7.95). A 2×2 ANOVA on the belief scores revealed the same patterns. The means were 67.53 and 52.46 for simultaneous lineups and 63.38 and 53.93 for sequential lineups. The interaction was nonsignificant, $F(1, 335) = 1.22$, $p = .27$, partial $\eta^2 < .01$, as was the main effect for lineup type, $F(1, 335) = 0.28$, $p = .60$, partial $\eta^2 < .01$. The effect for identification was significant, $F(1, 335) = 23.33$, $p < .001$, partial $\eta^2 = .07$, with a mean shift of over 10%.

Finally, the relationship between belief and verdict was examined. This is important as it is this final step in the decision process, deciding whether a defendant is guilty ''beyond a reasonable doubt''. The interest is whether the

[2] Applying Bonferroni's inequality for all 18 tests produces a critical pairwise α of $.05/18 = .003$. However, this is a very conservative test and arguably not all 18 tests are part of the same ''family'' of tests, and therefore α of .01 is a more appropriate conservative limit for pairwise comparisons.

relationship is similar for each condition. To examine this, a logistic regression was run to predict verdict using belief, experimental condition, and their interaction (see Agresti, 2002, for statistical details). The first thing to note is that there is the predicted large positive relationship between belief and verdict, $\chi^2(1) = 153.36$, $p < .001$, Nagelkerke $R^2 = .51$; for a 10% increase in belief in guilt, the odds of a guilty verdict increases by 2.85. The interaction between belief and condition was nonsignificant, $\chi^2(6) = 5.21$, $p = .52$, and there were no clear patterns in the parameter estimates to suggest further examination of the interactions. Therefore the interaction terms were removed. There was a main effect of condition, $\chi^2(6) = 16.64$, $p = .01$, which shows that after accounting for belief people in the different conditions still render different verdicts. There is a systematic pattern here. The two highest conditional probabilities are for the conditions where the suspect has been identified. If the data are reanalysed comparing these conditions with others, once controlling for belief, the odds ratio is 3.36 higher for suspect identifications than for the other conditions, $\chi^2(1) = 13.75$, $p < .001$. The level of belief corresponding to a 50% chance of a guilty verdict is about 62% for those in conditions where the suspect has been identified, compared with about 74% for the other conditions. From a legal perspective, these values can be interpreted as the reasonable doubt threshold for juror decision making (Wright & Hall, 2007). There were no other differences among conditions.

DISCUSSION

Previous research has shown that sequential lineups produce fewer false positives than simultaneous lineups. Although there also appears to be a decrease in the frequency of accurate identifications (Steblay et al., 2001), because of the relative utility of these two types of errors (i.e., Blackstone's adage that it is better that 10 guilty people are set free than 1 innocent person imprisoned), many psychology and legal experts have argued that sequential lineups should be used instead of simultaneous lineups. These findings have prompted several jurisdictions to begin using some forms of sequential lineup (e.g., in California, New Jersey, and New York). This means jurors will have to consider how to weight the value of eyewitness identification from these lineups. Here we examined this for both simultaneous and sequential lineups. Simultaneous and sequential lineups produced similar results with respect to verdict and to a lesser extent to belief in guilt. Positive identifications were associated with more guilty verdicts and higher beliefs in guilt when an identification was made compared with when a filler was identified and when no identification was made. The effect sizes were considerable, showing a large impact of eyewitness testimony. The proportion of guilty verdicts rose from about one quarter to over a half.

Identifying a filler and making no identification produced approximately the same levels of belief in guilt and numbers of guilty verdicts. In much of the police records these are combined and treated as "errors", but it is worth differentiating them. Wells and Olson (2002) showed that eyewitnesses choosing a filler often carries more diagnostic information than making no identification, and in certain circumstances may be more informative than even suspect identifications. This is because choosing a filler shows that the eyewitness is willing to choose somebody, and the person the eyewitness thinks looks the most like the culprit is not the suspect. Wells and Olson note that police often just assume someone choosing a filler is a "bad witness" and therefore discount the eyewitness as unreliable, rather than using this as exonerating evidence.

It appears that potential jurors give similar weight to an eyewitness identification when made in a simultaneous or a sequential lineup. It is worth examining the difference in diagnosticity between simultaneous and sequential lineups using the same measure, odds ratios, as reported here. Because there tend to be more target identifications and filler identifications in simultaneous lineups (Steblay et al., 2001), some researchers claim people just set a higher threshold for identifying someone in a sequential lineup than in a simultaneous lineup (Ebbesen & Flowe, 2003; Meissner et al., 2005). The largest comparison of simultaneous and sequential lineups to date comes from Steblay et al.'s meta-analysis (2001, their Table 1).[3] For simultaneous lineups, 50% chose the target in target-present lineups and 51% identified a filler in target-absent lineups. Two-thirds of the studies they analysed used six-person lineups, so an approximate estimate of

[3] Given the amount of recent interest in and research on sequential lineups, Steblay (personal communication, March 2006) and colleagues are planning a new meta-analysis on the topic.

the percentage of times an "innocent" target would be chosen in these is 8.5%. This produces an odds ratio of 10.76. For sequential lineups, 35% of the time the target was picked in target-present lineups. This suggests an estimated 4.7% of the time an innocent target is chosen. This produces an odds ratio of 11.0, which is very similar the ratio for simultaneous lineups. Ebbesen and Flowe reach a similar conclusion using a closely related technique (signal detection theory measures). Wells and Olson (2002) conclude that identifications with sequential procedure do provide more information about the guilt of the suspect, but they use a very different measure based on the ratio of the difference between probabilities, so more akin to the relative risk statistic (see Agresti, 2002, pp. 47, 124, for comparison, and Tredoux, 1999, for application to lineups). These values are all above those reported here, in Table 2. Of course the diagnostic value of an identification will depend on the value of other evidence.

The relationship between belief in guilt and verdict was examined for the different conditions using logistic regressions. This analysis was exploratory. Given that the number of guilty verdicts is small in some conditions (see Table 1) it is important to consider the overall pattern of results. Belief in guilt was, and should be, a very important variable for estimating the probability of rendering a guilty verdict. Beyond this, however, there was an effect by the lineup outcome. After statistically controlling for belief, the odds of rendering a guilty verdict were more than three times higher for people who were told that the suspect was identified, compared with the participants in the other conditions. To render a guilty verdict a juror is supposed to believe in guilt "beyond a reasonable doubt". This means people in the suspect-identified conditions used a different threshold for reasonable doubt from others. If deciding guilt simply involves having a numerical threshold of reasonable doubt and seeing whether your belief in guilt exceeds this threshold, then these data show that people's reasonable doubt threshold was about 10% lower for those in the suspect-identified conditions than in the other conditions. The participants were willing to render guilty verdicts at lower levels of belief in guilt.

What do these results mean for judicial practice? With the publication of *The Guidelines* and several jurisdictions using sequential lineups, the courts have to consider how jurors interpret eyewitness identifications from different types of lineups. Currently judges often warn jurors that eyewitness memory can be inaccurate when made with a simultaneous lineup. How this is done for sequential lineups is clearly important. It will be important to educate jurors about sequential lineups. They will have to be told about the procedure and also about the error rates for this procedure. The current research shows that people treat eyewitness identification from sequential lineups in a similar way to the way in which they treat identifications from simultaneous lineups.

Manuscript received 12 November 2005
Revised manuscript accepted 13 June 2007
First published online 13 July 2007

REFERENCES

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

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

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003.

Ebbesen, E. B., & Flowe, H. D. (2003). *Simultaneous v. sequential lineups: Factors preventing generalization of findings from the laboratory to the field*. Paper presented at the Psychology & Law: International, Interdisciplinary Conference, Edinburgh, 7–12 July.

Hilgard, E. R. (1953). *Introduction to psychology*. New York: Harcourt, Brace & Company.

Kebbell, M. R. (2000). The law concerning the conduct of lineups in England and Wales: How well does it satisfy the recommendations of the American Psychology-Law Society? *Law and Human Behavior*, *24*, 309–315.

Kraemer, H. C. (2004). Reconsidering the odds ratio as a measure of 2×2 association in a population. *Statistics in Medicine*, *23*, 257–270.

Levi, A. M. (1998). Are defendants guilty if they were chosen in a lineup? *Law and Human Behavior*, *22*, 389–407.

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

Loftus, E. F. (1974). Reconstructing memory: The incredible eyewitness. *Psychology Today*, *8*, 116–119.

Loftus, E. F., & Ketcham, K. (1991). *Witness for the defense: The accused, the eyewitness, and the expert who put memory on trial*. New York: St. Martin's Press.

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.

Penrod, S. (2003). Eyewitness identification evidence: How well are witnesses and police performing? *Criminal Justice Magazine*, *54* (Spring), 36–47.
Perrotta, T. (2004). *Police lineup rules questioned by judge: Says court can require ''double blind'' technique*. On http://www6.law.com/lawcom/ accessed 31 March 2004.
Scheck, B., Neufeld, P., & Dwyer, J. (2003). *Actual innocence: When justice goes wrong and how to make it right*. New York: New American Library.
Steblay, N. M., 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*, 459–474.
Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement (NCJ No. 178240)*. Washington, DC: US Department of Justice, Office of Justice Programs.
Tredoux, C. G. (1999). Statistical considerations when determining measures of lineup sise and lineup bias. *Applied Cognitive Psychology*, *13*, S9–S26.
Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, *17*, 969–993.
Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55*, 581–598.
Wells, G. L., & Olson, E. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied*, *8*, 155–167.
Wells, G. L., & Olson, E. (2003). Eyewitness identification. *Annual Review of Psychology*, *54*, 277–295.
Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling in social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*, 1115–1125.
Wilkinson, L., & the Task Force on Statistical Inference, APA Board of Scientific Affairs. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604.
Willing, R. (2002). Police lineups encourage wrong picks, experts say. *USA Today*, 26 November.
Wright, D. B. (2003). Making friends with your data: Improving how statistics are conducted and reported. *British Journal of Educational Psychology*, *73*, 123–136.
Wright, D. B., & Davies, G. M. (2007). Eyewitness testimony. In F. Durso, R. Nickerson, S. Dumais, S. Lewandosky, & T. J. Perfect (Eds.), *Handbook of applied cognition* (2nd ed., pp. 763–786). Chichester, UK: John Wiley & Sons.
Wright, D. B., & Hall, M. (2007). How a ''reasonable doubt'' instruction affects decisions of guilt. *Basic and Applied Social Psychology*, *29*, 85–92.
Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology*, *10*, 75–84.

Appendix

There were seven different conditions of this evidence. This is for condition 1.

Please Read Carefully

Seventeen-year-old Nancy Von Roper was hitch-hiking at about 6.30pm on the 12th of October, 1980. A light-blue compact car, driven by a bearded man, stopped, and she was offered a ride. After a few minutes he exited the motorway, before he should have for the location she thought that she was being taken. When she looked confused he said "I have to stop and see my sister". He drove some more and turned down a dirt road. He pulled to the side of the road. She felt an object at her throat. He said "Do as I say, or I'll hurt you". He had her undress and then raped her. He then had her get out of the car and he drove off.

She called the police at 7.22pm. She described the rapist as 25–30 years old, six foot, medium build, full beard and wearing a three-piece cream-coloured suit. The car was described as blue with temporary license plates displayed on the rear window. She remembered something hanging from the rear-view mirror and a brown folder on the back-seat. She showed police where the rape occurred. Tyre tracks were photographed. At 1.20am, police spotted a light-blue Chevette with temporary plates parked near to where the rape occurred. Tom Hoyle was in a restaurant with his fiancée, Gretchen Abraham. Hoyle was a manager of a restaurant chain, and drove the light-blue Chevette, a company car. Hoyle also fit the basic description given by Nancy Von Roper, except for being only 5′ 8″ (the victim never saw the person standing). The couple drove off and were soon pulled over by police.

The police asked Hoyle where he had been in the afternoon and evening. He explained he had been at his parents' house—it was his father's birthday—until 6.10pm. He went to his best friend's to watch TV at about 6.50pm and made some phone calls. Telephone records verified that there was a call made at 7.00pm from the house. Given the distance between his parents and his friend's, it would have been unlikely that he could have committed rape in this time. He left his friend's house to pick up Gretchen at 9.20pm. He gave police permission to search his car. He also gave them permission to take his photograph.

Tom Hoyle was arrested and his car impounded. He failed a lie detector test when asked if he had raped Nancy Von Roper. This can occur for a number of reasons, including the person telling the truth but being nervous. The victim stated that the number of the temporary license plate was either 667 or 776. The number on the license in Hoyle's car was six digits, 661-677. It is unlikely that there would be another automobile fitting the description with such a similar temporary license number. The police also reported that there was a folder like the one the victim described in the back of Hoyle's car.

Police took pictures of five other men, all with beards, between 25 and 30, and about the same build as Hoyle. On the 13th they showed these pictures with the one of Tom Hoyle to Nancy Von Roper. They showed all six photographs at the same time. Nancy Von Roper identified Tom Hoyle's photograph as the person who raped her.

In criminal trials, in order to make a verdict of guilty, you must believe in guilt beyond a reasonable doubt. Please tick appropriate for the verdict that you would give if you were a juror in this case and had this evidence presented.

☐ Guilty ☐ Not Guilty

Using a 0–100% scale, how strongly do you believe that Tom Hoyle is the person who raped Nancy Von Roper. 0% means that you are sure that he is not the rapist and 100% means that you are sure that he is the rapist.

Belief (write a percentage between 0–100%) ______