

By: Wendy Murphy, JD; Megan Mitchell, JD (expected 2013); Alexa Sardina, PhD (expected 2014)

Study Name and citation: Loftus, E. & Pickrell, J. E. (1995). The Formation of False Memories. *Psychiatric Annals*, 25(12), 720-725. (Commonly known as the “Lost in the Mall” study).

Theory Advanced by the Study: False childhood memories can be implanted into adults by repeatedly stating the false memory as true.

Conclusion: The study’s theory and methodology are not reliable under either *Frye* or *Daubert* standards.¹

Overarching Observation: Irrespective of scientific reliability, no study should be admitted as evidence in a court of law, whether directly or via the testimony of an expert relying on that study, unless it is relevant to an issue legitimately in dispute in a particular legal controversy. Relevancy standards applicable to scientific and “specialized” testimony and evidence refer to this as the “fitness” factor, which means the proposed evidence “fits” the facts because it will help to elucidate the truth on a disputed issue.

For example, any attempt to use the *Lost in the Mall* study to prove or disprove that a person is telling the truth about sexual abuse should be unsuccessful simply because memories about getting lost in a mall have nothing to do with memories of being sexually or physically violated. Put another way, the capacity to make a person falsely believe they were once lost in a mall does not “fit” and bears no relevance in a dispute over whether a person can be made to falsely believe they were sexually abused as a child.

This argument should be made *first*, before addressing more complicated scientific arguments about how the “*Lost in the Mall*” study was designed and conducted. It should be made clear to the judge that if a particular study, or proposed expert testimony relying on that study, is not relevant because it does not “fit” with the facts in dispute, there is no need to undertake a burdensome analysis of whether scientific standards were satisfied during the research process. Simply put, if it isn’t relevant, it doesn’t matter that a study is scientifically unreliable.

IMPORTANT DISCLAIMER: *Daubert* standards apply in federal court, and in some, but not all, state courts. Most states apply some version of *Daubert* or *Frye*, though several have their own criteria, usually a variation on *Daubert* or *Frye*. This analysis is meant only as general guidance on legal principles involving the admissibility of specialized/scientific evidence. It is the opinion of the authors and is not intended as legal advice for any particular case or as a statement of law in any particular jurisdiction. Litigants should consult with an attorney to determine whether the analysis here is applicable and sufficient in their jurisdiction.

¹ It is critically important to determine the applicable standards in a particular jurisdiction before submitting an analysis in court. Every state has a rule or case that explains how a *Daubert* or *Frye* analysis works in a particular jurisdiction. A simple internet search for “[STATE] admissibility of scientific evidence” should provide ample information about the factors to be applied.

COPYRIGHT INFORMATION: This analysis is copy written by Wendy J. Murphy, JD (2013) and cannot be used without attribution and permission of the author. All rights reserved.

1. DAUBERT ANALYSIS

To be admissible as evidence in a court of law, a study must be analyzed to determine its methodological soundness and reliability. This is done by examining the following factors: A. whether a study is capable of being tested and replicated; B. whether a study was subject to peer review and publication; C. what the known or potential error rate is; D. the existence and maintenance of standards and controls; and E. the degree to which a study has been generally accepted.

A. Empirical testing and replication

A study is reliable under this factor if it has been tested and the results can be replicated. If the study has not been tested or is incapable of being tested, then it cannot meet the requirements of this factor. While the *Lost in the Mall* study is capable of being tested, its results have never been replicated. Replication is the process of repeating a study using the same methods by different subjects and researchers.² A study by Pezdek, et al., at the University of Claremont in 1997 comes close to replication, but lacks similar enough criteria to be considered a true replication.

Like *Lost in the Mall*, Pezdek et al., used relatives as study subjects, however, the relatives in *Lost in the Mall* were used only to provide true memories about family shopping trips. Pezdek et al., used the relatives to implant false memories in their younger relatives.

While other studies come close to replicating the methods used in *Lost in the Mall*, the results themselves have never been replicated. This is important because replication of **results** is what makes the methodology and statistical significance of the original experiment reliable.³ *Lost in the Mall* fails the reliability test under this factor for two reasons. (1) It has not been repeated using the same methodology⁴ and (2) similar studies have obtained widely varied results.

When asked about problems with replication and a high error rate (addressed more specifically below) in 2007, the lead author of *Lost in the Mall*, Elizabeth Loftus, stated “First of all, numerous other researchers have gone on to adopt this methodology and they get much higher rates of subjects falling for the suggestion so I don’t have to defend the 25 percent rate when other people, I mean, are getting 3 percent or 50 percent false memory rates in these

² <http://www.experiment-resources.com/validity-and-reliability.html>

³ *Id.*

⁴ *Id.*

studies.”⁵ In other words, she claims she does not have to defend a high error rate in her work because other studies using the same procedures got lower error rates.⁶

This explanation makes no sense because replication looks to the specific conditions of the study and whether, under those conditions, the study results can be replicated. While other researchers may have adopted similar procedures to those used in *Lost in the Mall*, the study results themselves have not actually been replicated.⁷

B. Whether the study was subjected to peer review and publication

A study is reliable under this factor if it has been subject to peer review and publication. *Lost in the Mall* was published in *Psychiatric Annals*, a journal ranked 78th in the world for 1999 according to SCImago Journal and Country rank.⁸ Being published in a journal, especially if the journal lacks an effective peer review process, carries little weight⁹ because the quality of the peer review process, more than the fact of publication, determines whether the study has been assessed *effectively* by objective reviewers.

The public record reveals little about whether *Psychiatric Annals* is in fact a peer-reviewed journal. Without more information about the nature of the review process employed in the decision to publish *Lost in the Mall*, it is difficult to determine the study’s reliability under this factor.

C. Known or potential error rate

A study is reliable under this factor if there is a known or potential error rate *and* the known error rate is low enough to substantially reduce the possibility that the results were reached due to an accident or fluke. *Lost in the Mall* does not report any actually calculated error rate, which is a serious flaw because error rates are critically important to a fair assessment of a study’s reliability.¹⁰ An error rate indicates the amount of variability among results such that a researcher will discover the likelihood that a mistake has occurred. Ideally, an error rate should be as small as possible, however a statistically insignificant rate is sufficient to support a finding of reliability. An error rate should be .05 or lower. When a study crosses this threshold it becomes more likely that human error caused the results, which increases the risk that a study’s results are unreliable.¹¹ As noted below, an apparent error rate in *Lost in the Mall* was reported

⁵ Liano v, The Diocese of Phoenix, (pp.211-212).

⁶ [http://faculty.washington.edu/eloftus/Articles/AmerPsychAward+ArticlePDF03%20\(2\).pdf](http://faculty.washington.edu/eloftus/Articles/AmerPsychAward+ArticlePDF03%20(2).pdf)

⁷ **Replication** is the repetition of an earlier experiment to duplicate (and perhaps extend) its findings. **Direct replication** is repeating an experiment as closely as possible to determine whether the same results will be obtained. **Systematic replication** is the repetition of an experiment while varying numerous factors, such as the use of relatives as informants, to demonstrate the lack of significance of those factors by establishing that a study’s findings survive those changes.

⁸ <http://www.scimagojr.com/journalsearch.php?q=18684&tip=sid>

⁹ Mahoney (1977) “Publication Prejudices”, Peters & Ceci (1982) “Peer Review Processes of Psychological Journals”, and Mahoney et al. (1978) “Getting Published”

¹⁰ See Elmes, D. G., Kantowitz, B. H., & Roediger, H. L. (2012) “Research Methods in Psychology”. United States: Wadsworth Cengage Learning.

¹¹ <http://homepage.psy.utexas.edu/homepage/class/Psy391P/Josephs%20PDF%20files/Schmidt.PDF>

twice in terms of the number of subjects that did not have a false memory implanted, but the rate was never actually calculated and published.

The fact that an error rate was never calculated indicates an attempt to avoid publishing unwanted results.¹² An attempt to account for this glaring anomaly is apparent because the author writes about “clarity ratings” as a seeming alternative to an error rate. This raises even greater concerns and adds nothing to an assessment of reliability because “clarity rating” is a meaningless term under *Daubert*.

While the author of *Lost in the Mall* calculated no error rate herself, the error rate that can be gleaned from the published data appears to be substantial.

Lost in the Mall reported two sets of results. First, it was reported to the HSC that 2 out of the 24 subjects believed the false memories to be true in 1994, and then it was reported in 1995 that 6 out of 24 subjects believed the false memories to be true.¹³ The change from 2 out of 24 to 6 out of 24 suggests a high rate of error.

Indeed, this amounts to a 9.091% and a 33.33% rate of error respectively.¹⁴ The acceptable rate of error for a study to be deemed methodologically reliable is 5%.¹⁵ The author’s failure to calculate and report this weakness in the data is highly unorthodox.¹⁶

The author’s use of “clarity ratings” further undermines the study’s reliability under this factor for two reasons: (1) “clarity ratings” are not an accepted technique for proving the validity of a study; and (2) even if “clarity ratings” were acceptable, the author relied on only narrowly described segments of memories to determine “clarity” based on an assessment of falseness to determine “clarity.” This is inadequate simply because a small bit of information about a partial memory is not comparable to a full memory for an event or narrative scene.

D. The existence and maintenance of standards and controls

Under this factor a study is reliable if standards and controls were used to prevent unaccounted for variables from undermining the reliability of the results. *Lost in the Mall* is not reliable under this factor because of insufficiently maintained standards and controls, including but not limited to the following:

(1) The study had a sample size of only 24 subjects. A small sample size coupled with a relatively high error rate produces a statistically insignificant result.¹⁷ A large error rate can be

¹² Chow, S. L. (2002). STATISTICS AND ITS ROLE IN PSYCHOLOGICAL RESEARCH. In *Methods in Psychological Research*, In *Encyclopedia of Life Support Systems (EOLSS)*, Eolss Publishers, Oxford, UK, [<http://www.eolss.net>]

¹³ See Elmes, D. G., Kantowitz, B. H., & Roediger, H. L. (2012) “Research Methods in Psychology.” United States: Wadsworth Cengage Learning.

¹⁴ Calculations of the rate of error were completed using a percent rate of error calculator found at: <http://www.calculator.net/percent-error-calculator.html>

¹⁵ <http://www.psych.umn.edu/courses/spring05/mcguem/psy8935/readings/lykken1968.pdf>

¹⁶ *Supra*, note 13.

¹⁷ <http://core.ecu.edu/psyc/wuenschk/stathelp/Type1.htm>

addressed by re-conducting the experiment with a larger sample size,¹⁸ but that was not done here. Such a small sample size, irrespective of error rate, renders the results largely unreliable.¹⁹ A small sample size is even more problematic when the sample consists of an unevenly mixed group of diverse subjects such that the results cannot reliably be replicated.²⁰

(2) The study was not performed under conditions of uniformity. This is demonstrated in several ways:

a. It failed to use a control group. Without a control group it is impossible to demonstrate the reliability of the results because the data fail to illustrate whether the retention of memories would have been similar for similar people under similar circumstances. This makes it possible that unaccounted for variables are actually controlling the outcome of the study.

b. It included wide diversity in age of subjects (18-53), which undermines uniformity and reliability because frontal lobe development is not even complete until between the ages of 20-22. Therefore, the subjects within the study group are not sufficiently similar to one another to produce reliable results. (Bruck & Ceci, 1999).

c. It included only 3 males. This undermines uniformity and reliability because of the overrepresentation of females.²¹ Other studies have shown that men and women perceive situations differently, yet the author nowhere explains or accounts for this difference.

d. It failed to ensure that study participants were relating memories within the same relationship of trust (e.g., all older relatives). Failure to use the same relationship across and within groups undermines uniformity and consistency. (Pasakoff et al. (2000)).

e. It failed to utilize “item correlation” criteria to ensure that the instrument used (“the mall story”) measured the same construct utilized in the study. An item-total correlation test is performed to determine whether any item in the set of tests is inconsistent with the averaged behavior of the others, thus can be discarded. The analysis should have been performed to eliminate any inconsistent items prior to determining the factors that represent the construct, in this case “the mall story.”

f. It failed to account for memory failure based on age, such as why it might be more difficult for a 53 year-old to remember being lost in the mall 48 years earlier as compared to an 18 year-old remembering being lost in the mall 13 years earlier. This undermines the study’s reliability because other research has shown that memory is affected by the passage of time. Study subjects should have been roughly the same age to comply with the reliability requirement of consistency. (Mitchell, Johnson, and Mather, 2003).²²

¹⁸ <http://people.musc.edu/~elg26/teaching/psstats1.2006/maccallumetal.pdf>

¹⁹ *Id.*

²⁰ See Elmes, D. G., Kantowitz, B. H., & Roediger, H. L. (2012)“Research Methods in Psychology”. United States: Wadsworth Cengage Learning.

²¹ Out of the 24 subjects only three were male.

²² Mitchell, K.J., Johnson, M. K., & Mather, M. (2003). Source monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology*, 17, 107-119.

(3) The study instructions were provided to subjects by two researchers who also coded the data. However, the study provides no information pertaining to inter-rater reliability (r squared). The inter-rater reliability measurement is used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon. Without this measurement, the study loses inherent reliability (Elmes et al. 2012). Similarly, the study involved the use of family members to provide the true memories, which were used to implant the false memories. Use of relatives raises serious concerns about reliability of the results because family members are inherently biased. The study also fails to indicate whether the subjects were properly informed about objective standards for communicating substantive information about memories to the other participants. This is critical to maintain internal consistency and objectivity (Elmes et al. 2012).

(4) No attempt was made to look for discrepancies among gender, age group, and the specific age being used in terms of the age of the subject when the false memory was supposed to have happened. With such a small sample size, failure to factor for variations in characteristics of subjects involved in the study further diminishes reliability (Elmes et al. 2012).

While some studies are designed with weaknesses that cannot be avoided, it is imperative that researchers reveal these weaknesses and explain the reasons why they may have had a significant effect on the outcome. *Lost in the Mall* offers no justification for the presence of many weaknesses in reliability and scientific integrity.

Without sufficient maintenance of standards and controls, there is simply no way of knowing why a study subject might have believed he or she had been lost in a mall as a child, and it cannot be said that the belief was false or implanted. Indeed, under the woefully inadequate controls applied here, it cannot even be determined whether those who denied having such a memory might have falsely forgotten a true event.

E. The degree to which the theory and technique are generally accepted by the relevant scientific community

To be reliable under this factor, the theory or technique at issue regarding a study must be generally accepted by its relevant scientific community. (See also Frye analysis in Section II below). Even if the theory advanced by the study is not “generally” accepted, the degree of acceptance indicates whether the study is more or less reliable. The techniques used in *Lost in the Mall*, as outlined in section D above, lack sufficient indicia of reliability to merit general acceptance by the relevant scientific community.

In addition, the relevant scientific community would not likely accept this study as a reliable assessment of how memories work in general, or whether false memories can be implanted, because to be considered generalizable and acceptable by the psychology community, the results must be replicated, which has never been done with *Lost in the Mall*.²³

²³ <http://www.experiment-resources.com/validity-and-reliability.html>

II. FRYE ANALYSIS

To be admissible as reliable evidence under *Frye*, the theory advanced by the study must be generally accepted by its relevant scientific community. To gain general acceptance in the scientific community a study need not be free of criticism, rather, it must be deemed valid and reliable by a significant segment of the relevant scientific community. When determining whether a study is generally accepted, courts consider criticisms of the study, as well as information in manuals, encyclopedias or other publications that are respected or endorsed by its relevant scientific community.²⁴

While *Lost in the Mall* makes no claim regarding theories pertaining to children and adults regarding suggestibility and memory, there have been many articles, as well as attempts by lawyers and expert witnesses in legal proceedings, suggesting that *Lost in the Mall* supports the idea that false memories of abuse can be implanted in children and adults. To the extent “acceptance” can be measured by the amount and content of those articles, it is fair to infer that the study’s theories and conclusions have not been generally accepted because a majority of commentary rates the study as unreliable and lacking in sound methodology.

In Crook and Dean’s (1999)²⁵ critique, for example, *Lost in the Mall* is described as lacking in support by the relevant scientific community simply because the study had not been examined to ensure that the research was conducted properly and reported accurately.

Others have said that *Lost in the Mall* lacks general acceptance simply because, although it has been demonstrated that a person can be made to believe falsely that a past event occurred, (Hyman, Husband, & Billings, 1995)²⁶ such studies are not generalizable in terms of supporting claims about all *types* of memories. Memories such as being lost in a mall are commonplace, relatively benign and plausible. Memories of significantly different events, such as sexual abuse, are not similar enough to memories of being lost in a mall such that conclusions about the ability to implant one type of memory allow conclusions to be drawn about the ability to implant the other. Indeed, studies that attempted to implant false memories more akin to sexual abuse resulted in a success rate of zero. For instance, Pezdek and Roe (1997)²⁷ attempted to implant false memories of participants receiving rectal enemas as children. However, they reported 0% success rate, whereas they were able to persuade 15% of study subjects to believe, falsely, that they had once been lost in a mall.

Overall, though a number of studies have demonstrated that false memories can occur, the data clearly shows that the *type* of event falsely recalled is significant and that being made to falsely recall getting lost in a mall has no bearing on whether a person can be made to falsely recall something as different and significant as sexual abuse. (Berliner & McDougall, 1997; Porter & Marxsen, 1998)²⁸. Because the possibility that people can be made to develop false

²⁴ For example some courts have seen publication of a disorder in the DSM as being per se generally accepted.

²⁵ Crook, L. S. & Dean, M. C. (1999). “Lost in a shopping mall” – A breach of professional ethics. *Ethics & Behavior*, 9(1), 39-50.

²⁶ Hyman, I. E., Husband, T. H., & Billings, J. F. (1995). False memories and childhood experiences. *Applied Cognitive Psychology*, 9, 181-197.

²⁷ Pezdek, K. & Roe, C. (1997). The suggestibility of children’s memory for being touched: Planting, erasing, and changing memories. *Law and Human Behavior*, 21, 96-106.

²⁸ Berliner, L. & McDougall, J. (1997). Agenda for research: Clinical approaches to recollections of trauma. In J. D. Read & D. S. Lindsay (Eds.), *Recollections of trauma: Scientific evidence and clinical practice* (pp.523-530). New York: Plenum.

memories for traumatic incidents similar to sexual abuse has not been reliably studied, much less demonstrated, it cannot be said that *Lost in the Mall*, to the extent it supports the theory that people can be made to falsely recall sexual abuse, has been generally accepted by the relevant scientific community. (Porter, Yuille, & Lehman, 1999)²⁹.

SOURCES OF CRITICISM

Crook, L.S., & Dean, M. C. (1999). "Lost in a Shopping Mall" – A breach of professional ethics. *Ethics & Behavior*, 9(1), 39-50.

Bruck, M. & Ceci, S. J. (1999). Suggestibility of children's memory. *Annual Review of Psychology*, 50(4), 419-439.

Elmes, D.G., Kantowitz, B.H., & Roediger, H.L. (2011). *Research methods in psychology* (9th Edition). Belmont, CA: Wadsworth Publishing.

Haaga, D.A. (1993). Peer review of term papers in graduate psychology courses. *Teaching of Psychology*, 93, 28-32.

Smith, N.C. (1970). Replication studies: A neglected aspect of psychological research. *American Psychologist*, 25(10), 970-975.

Further Readings

Norman, K.A. & Schacter, D.L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memory. *Memory and Cognition*, 25, 838-848.

Pezdek, K. K. & Roe, C. (1997). The suggestibility of children's memory for being touched: Planting, erasing, and changing memories. *Law and Human Behavior*, 21(1), 95-106.

Pilemer, A.T. & White, S. H. (1988). Very long term memories of the first year in college. *Learning, Memory, and Cognition*, 14, 709-715.

Podsakoff, P.M., MacKenzie, S.B., Yeon-Lee, J. & Podsakoff, N. P. (2000). Common method biases in behavioral research: A critical review of the literature and recommended studies. *Journal of Applied Psychology*, 88(5), 879-905.

Porter, S. & Marxsen, D. (1998). Challenging eyewitness expert: An update considering memory for trauma and false memories. In J. Ziskin (Ed.), *1997 Supplement to coping with psychiatric and psychological testimony* (5th ed., pp. 112-129). Los Angeles: Law and Psychology Press.

²⁹ Porter, S., Yuille, J. C., & Foster, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior*, 23, 517-537.