MEMORY, 2008, 16 (4), 436–442

Psychology Press
Taylor & Francis Group

The co-witness misinformation effect: Memory blends or memory compliance?

Elin M. Skagerberg
University of Sussex, Brighton, UK

Daniel B. Wright
University of Sussex, Brighton, UK, and Florida International University, FL, USA

For 30 years hundreds of researchers have shown participants videos and slide-sequences of events, presented the participants with misleading information, and found that this misinformation distorted their memories. The purpose of this study was to establish whether those misled participants are reporting a memory blend of the two sources of information or whether they are simply complying with the post-event information. A total of 92 participants were shown one of two versions of six different videos, which included some subtle differences. After having watched each video individually, participants were paired with someone who had seen the other version and they discussed the clips together. They then individually answered questions about the videos, and their responses showed that some of the distorted memories were blends of the original information and the post-event information. The implications of these findings are discussed.

Memory is malleable (Loftus, 2005). It is a reconstructive process that can be influenced by a person's general knowledge of the world (Norman, Rumelhart, & The LNR Research Group, 1975) and by post-event information (Wright & Loftus, 1998). This post-event information (PEI) can be encountered non-socially (e.g., Lindsay, 1990; Loftus & Palmer, 1974) or socially (e.g., Gabbert, Memon, Allan, & Wright, 2004; Wright, Self, & Justice, 2000). PEI research has allowed theoreticians to focus on how people reconcile contradictory information in memory and has allowed practitioners, particularly within eyewitness research, to assess the potentially detrimental effects when eyewitnesses encounter misleading PEI (Loftus, 2004). The critical question is about the nature of an event memory that has been produced by showing people both the original event and the PEI. Research shows that some people report the original information and some people report the PEI, but is this either due to people's memory being mainly of one of the two events or due to the memory being some blend of the two? Or is it possible that these answers simply reflect strategic compromise responses?

The stimuli used in most PEI research involve showing an original object, for example a Stop sign, and then suggesting in the PEI that it was something else, such as a Yield sign. This choice of stimuli does not allow researchers to tell if participants can have some blend memory because people will report that they saw some recognisable object, and there is not a continuum of equally recognisable objects between a Stop and a Yield sign that could correspond to a blended memory. Thus people's memory responses would be one of the two traffic signs even if their actual memory representations corresponded to a blend of the two (Metcalfe,

Address correspondence to: Elin M. Skagerberg, Psychology Department, University of Sussex, Brighton, BN1 9QH, UK. E-mail: elins@sussex.ac.uk

http://www.psypress.com/memory DOI:10.1080/09658210802019696

1990; Schooler & Tanaka, 1991). In order to investigate the presence of blend memories it is necessary to use stimuli where blends are possible, and these are likely to be those that have several intervening alternatives between the original information and the PEI.

Within the traditional non-social PEI research, several studies have used stimuli that could produce memory blends, such as speed (Loftus & Palmer, 1974), time (Burt, 1999; Burt & Popple, 1996), colour (Belli, 1988; Loftus, 1977), and the number of people (Loftus, 1975). For example, Loftus (1977) showed participants a car accident involving a green car and later suggested that it was blue. The participants were asked to indicate the colour of the car from 15 patches and Loftus found that the control participants opted for patches in the green range, although with some variability, while the distribution of those presented with misinformation was shifted towards the blue patches. The critical question is whether the shift was due to some participants responding as if they saw green and some responding as if they saw blue, or whether some of the responses were based on a blend of the two colours. In other words, can the distribution from the PEI group be accounted for by adding together the green and blue distributions? We describe these two possibilities as memory compliance and memory blending, respectively. It is possible that within any study there may be some participants who show compliance and some who show blending. Because some participants are likely not to process either the original or the post-event information, some will be responding like those only shown one of the stimuli. Therefore, the purpose of this study is to show whether memory compliance can account for all of the PEI responses.

In Loftus (1977) the modal response of those presented with PEI was between green and blue, which suggests that some people have blend memories, but this is not necessarily the case. The top panel of Figure 1 shows a normally distributed curve centred on the blue part of a colour dimension; the second panel shows a normally distributed curve centred on the green part of this dimension; and the final panel shows a curve that is based on half of the people being from the blue distribution and half from the green distribution. As can be seen, the resulting curve is uni-modal and centred in between blue and green. While intuitively a uni-model distribution may suggest blend memories, in fact the distribution in the final panel is simply the sum of the distributions in the top two panels. Thus, a uni-model distribution does not imply that blend memories exist (see also Waller & Meehl, 1998, ch. 2).

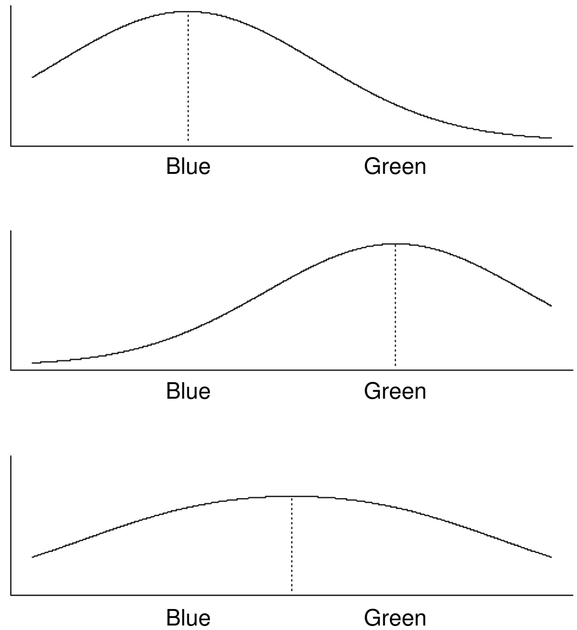

Figure 1. The top panel shows responses centred on blue, the middle panel on green, and the final pattern is 50% from the first distribution and 50% from the second distribution, and this creates a uni-modal distribution centred between blue and green.

In order to test whether the distribution of those who encounter PEI can be accounted for by assuming that some of the people are like those only shown one of the stimuli and the remainder are like those shown the other stimulus, we produce mixtures of different proportions from the observed distributions of those people who just saw one of the two stimuli and see whether these mixture distributions are significantly different from the PEI distribution. The traditional methods for this (for example, Macdonald & Green, 1988) force the user to assume a specific shape for the different distributions and the resulting statistics are sensitive to variations from these shapes. Therefore a significant effect can be due to the distributional assumptions not being met. An alternative is to make no distributional assumptions and compare different distributions using a non-specific test like the Kolmogorov-Smirnov test (for algorithms see Wright & Skagerberg, 2008). Here we test the more specific hypothesis whether any weighted combination of the control groups can account for the middle of the experimental group's distribution. For each possible proportion we calculate the chi-squared value for recreating the proportions in the different thirds of the PEI groups. If the minimum chi-squared value is still statistically significant, then this shows that we cannot recreate the observed distribution based on memory compliance, and therefore it shows that memory blending has probably occurred (details of the statistical procedure in Wright & Skagerberg, 2008).

The memory blending explanation described in Loftus (1977) is based on the trace alteration hypothesis. According to this theory the memory of the original event is often distorted and changed by misinformation. Holliday, Douglas, and Hayes (1999), for example, manipulated memory trace strength in 5- and 9-year-old children. The children were presented with a story and the next day were presented with a summary containing misinformation. The trace strength was manipulated using repetition. The findings showed that suggestibility levels were highest when the original trace was weak and the post-event trace was strong. This shows that misinformation degrades memory and that this memory alteration is dependent on the combined trace strength of both the original information and the misinformation (Holliday et al., 1999).

Recently there has been a move to use more social and ecologically valid means of presenting PEI to participants (e.g., Gabbert, Memon, & Wright, 2006; Mori, 2007; Paterson & Kemp, 2006b; Skagerberg & Wright, 2008). There are often several eyewitnesses to a crime and these people usually discuss aspects of the crime (Paterson & Kemp, 2006a; Skagerberg & Wright, in press). The most likely way in which eyewitnesses encounter PEI is by another eyewitness, and because memory can be inaccurate and, as the research cited above shows, inaccurate information can be passed from one eyewitness to another, this is one of the reasons why errant eyewitness memory is the leading cause of false conviction (Doyle, 2005).

We use co-witnesses to present errant information about stimuli that vary along a continuum in such a way that viable intermediate alternatives exist, allowing us to examine the blend hypothesis with socially presented PEI. The variables we use are: colours, number of people, weather, and the height and build of a person. The study is important for police and legal purposes, such as judging eyewitness reliability, and for other purposes, like insurance claims (Loftus, 1975).

METHOD

Participants

A total of 92 participants were recruited from the University of Sussex participant pool and received either £2.50 ($\sim$\$5) or course credits. Of these participants, 50 were placed in the pair condition (Mean age = 22.00 years, $SD = 4.34$; 64% female) and 42 in the control condition (Mean age = 21.98, years, $SD = 5.21$, 64% female). One participant in each pair was randomly allocated to see version A of the film clips and the other person saw version B. Those in the control condition were randomly allocated to see version A or B with the restriction that there were equal numbers in the two conditions.

Materials

The participants were shown six 5-second film clips, which were presented on PowerPoint. Three of the clips (numbers 2, 4, 6) had two versions with slight differences. We refer to these as the critical clips. In the first critical clip (number 2), version A showed a woman wearing a green scarf and version B showed the same woman wearing a light blue scarf. From the colour chart in Paint (from Define Custom Colours), their hues were 80 (green) and 172 (blue) on the non-normalised HSB/HSV scale, which goes from 0 to 360 degrees (from red to red), respectively. In clip number 4, version A showed a dark-haired, slender, tall man stealing a mobile phone while in version B the culprit had light-brown hair, and was shorter and stockier. Both men were White and in their mid-20s. In version A of the third critical clip (number 6) 18 people were shown waiting at a train station while in version B 6 people were shown waiting for the train. In addition, in the clip with 18 people the weather was slightly grey whereas in the clip with 6 people the weather was very grey. All other details of the clips were identical. The differences between clips all represent stimuli that vary on a continuum and for each clip there are viable intermediary possibilities. The response anchors were chosen to allow accurate responses for both versions of the film, but also to represent the range of possible answers for that type of question.

There were six questions about the differences between the video clips. The responses were made on 7-point rating scales, by free recall, or in the case of the scarf's colour using the colour chart in Paint. The six critical questions were:

- Critical question 1 (clip 2): What was the colour of the woman's scarf? Please write down the number corresponding to the colour that best describes the scarf using the colour chart provided.
- Critical question 2 (clip 4): How dark was the mugger's hair? (1 = very light, 7 = very dark)
- Critical question 3 (clip 4): Approximately how tall was the mugger? (responses were allowed in feet and inches, or centimetres)
- Critical question 4 (clip 4): Please indicate with a cross on the line the build that most applies to the mugger. (0 = very thin, 10 = very stocky)
- Critical question 5 (clip 6): How many people did you see in the clip?
- Critical question 6 (clip 6): What was the weather like? (1 = very grey, 7 = very sunny)

Procedure

Pair condition. Two previously unacquainted participants were asked to come to the laboratory at the same time. One was randomly allocated to see version A of the clips and the other to see version B. They each sat in front of a computer, with a separation wall between them so that they could not see each other's computer. They were told that the reason they were using separate computers was to ensure that they sat at the same distance from the screen. They were told to watch the clips carefully and were shown the first clip (non-critical). The pair were then directed to a table and told to discuss the clip. On the table was a sheet that included some basic guidelines of what to discuss. For the critical clips these directed participants towards the critical differences (e.g., colours, number of people, and location). The discussions evolved differently in the various pairs, with participants sometimes arguing and sometimes agreeing immediately. Following the discussion each person went back to their individual computer and answered three questions for the clip. When both participants had finished they were shown the second clip and the same procedure was used for the remaining clips. The critical clips were the second, fourth, and sixth clips.

Those in the individual conditions were randomly allocated to watch either version A or B, and then proceeded as those in the pair condition, but with an unrelated filler task rather than the discussion at the end of the session. Participants were debriefed, thanked, and paid/given course credits for their participation. Ethical clearance for this study was given by the Life Science Research Governance Committee and APA guidelines were adhered to throughout.

RESULTS

This section is divided into two parts. First, we identify which items show misinformation effects by conducting 2×2 between-group ANOVAs on each of the critical questions. Second, where large enough misinformation effects occur, we examine whether the distribution of the PEI group can be accounted for by a weighted average of the distributions of the two control condition distributions.

Table 1 shows the means and the standard deviations for the critical questions for the four conditions. The dependent variables were the questions (e.g., colour, height, number of people) and the independent variables were the version of the clip (A or B) and the condition (pair or control). The prediction for the misinformation effect is that the means for people in the pair conditions for versions A and B should be between the means for the control participants shown versions A and B. For example, the mean for the reported number of people seen at the train station in clip number 6 in the control condition was 16.71 ($SD = 4.27$) for group A (true value = 18) and 6.09 ($SD = 1.09$) for group B (true value = 6), while the mean reported number of people seen in the pairs condition was 12.08 ($SD = 4.47$) for group A and 6.72 ($SD = 1.46$) for group B. The expectation is that the means for the pair conditions in Table 1 should be between the values for the control conditions. The chance probability of having any mean between the two controls is 33%. As shown in Table 1, 10 of the 12 means are between these values.

ANOVAs (2×2) were run for all six critical questions. The interactions measure the misinformation effects and are shown in Table 1 with their corresponding effect size and p value. Except for the culprit's build, all the interaction effects are in the predicted direction. Three of the interactions,

TABLE 1
Means and SD: Pairs and control conditions

	Pairs		*Control*		*Interaction*
Question	*A*	*B*	*A*	*B*	$F(1, 88)$, η^2_p, p
Hair					
A = Medium brown	4.76	5.20	4.57	5.66	1.61, .02, .21
B = Very dark brown	(1.27)	(1.25)	(1.36)	(1.02)	
Height					
A = 177cm	178	179	176	179	0.33, .00, .57
B = 189cm	(5.07)	(9.24)	(4.02)	(4.31)	
Build					
A = stocky	5.80	5.05	5.13	4.60	0.10, .00, .75
B = thin	(1.59)	(1.77)	(1.52)	(1.66)	
People					
A = 18	12.08	6.72	16.71	6.09	15.11, .15, <.001
B = 6	(4.47)	(1.46)	(4.27)	(1.09)	
Weather					
A = slightly grey	2.72	2.88	3.38	2.52	8.98, .09, .003
B = very grey	(0.54)	(1.01)	(0.86)	(0.75)	
Colour					
A = 172	149.72	112.88	174.19	79.81	6.86, .07, .01
B = 80	(49.54)	(68.31)	(40.89)	(43.68)	

The means and standard deviations (in parentheses) for participants in the pairs and control conditions divided by whether they saw video version A or B. The statistics for the interaction of the 2 × 2 ANOVAs are shown in the final column.

for the number of people at the station, the weather, and the colour of the scarf, are substantially larger than the others (none of the others approach traditional levels of statistical significance). Wright and Skagerberg (2008), Fig. 7d) conducted several simulations and found that in order to have a good opportunity to detect a non-mixture, while still controlling for Type 1 error, the mean should be about halfway between the two control groups. Further, they found that the test only has adequate power if the control groups are separated by about three standard deviations or more. Only one of the conditions meets these standards. Therefore we examine the condition where people saw 18 people, but talked with someone who saw 6 people. The statistical function described in Wright and Skagerberg (2008) showed that the mixture of 30% from the first control (saw 6) and 70% (saw 18) of the control distribution was the closest for approximating the experimental distribution, but that it did not capture the middle of the distribution adequately—$\chi^2(2) = 7.50$; bootstrap adjusted $p =$ .02. Thus we can reject the idea that this observed distribution arose from some people behaving as though they had just seen the 18 people or were responding like those who had just seen 6.

DISCUSSION

There are two main findings. First, when pairs of participants who saw one of two versions of a video clip were asked to discuss the clip, their memories/answers were affected. Second, the means for these pairs were mostly in between the means for the control conditions shown these versions but without the discussion. While this suggests that memory blends may have occurred, more thorough statistical analyses are necessary for this claim (Figure 1). When these procedures were used they showed that the responses people gave individually after the discussion could not be modelled by saying that some of the people were simply choosing the misinformation. Instead, the analysis showed that memory blends occurred.

The method used to present PEI here is more ecologically valid than in traditional PEI research because eyewitnesses often claim to talk with other eyewitnesses. While several researchers have shown that this method can produce misinformation effects, possibly even larger misinformation effects than the traditional methods (Gabbert et al., 2004), it is important to show that this method can reproduce misinformation effects with a wide variety of stimuli, for example with the number of people remembered, the colour of a scarf, and the weather. Of course, further moves towards recreating the situations that typical eyewitnesses encounter and testing more types of stimuli are important to increase ecological validity and generalisability.

The fact that memories of eyewitness details could be easily manipulated by misinformation from a co-witness is worrying, as eyewitnesses often discuss the criminal event (Paterson & Kemp, 2006a; Skagerberg & Wright, in press) and as co-witness discussion has a particularly strong influence on eyewitness memory (Paterson & Kemp, 2006b). Examples such as the 1995 Oklahoma bombing, where one person's apparently errant memory led two others to have memory for an accomplice with McVeigh (Memon & Wright, 1999; Schacter, 2001), and the murder of the Swedish prime minister, where the witnesses were put in the same room during the investigation (Granhag, Ask, & Rebelius, 2005), show that such discussions can lead the police in the wrong direction. This could be time-consuming for the police, lead to miscarriages of justice for the innocent, and allow the guilty to remain at large.

The critical question of PEI research is about the memory of those who have encountered misinformation and this was the central focus of the current study. We described memory compliance as the process where a person could either respond with what they had originally seen or with what the other person saw. We adapted an existing statistical function for comparing mixtures of distributions and found that memory compliance could not account for the observed distribution. This shows that blend memories are reported. Although these results indicate that memory blends occur, they do not differentiate between conscious compromise responses and composite representations.

Yet the current findings are important as they show that a co-witness's memory of an event can blend with a witness's own memory of that event following discussion. Considering how often co-witness discussions occur (Paterson & Kemp, 2006a; Skagerberg & Wright, in press), it is important for juries and law professionals to keep the possibility of memory blends in mind. Law professionals should be aware that a witness's memory of an event, in court, could be a blend of his/her original memory and that of a co-witness. So why do memory blends occur?

A large number of studies show that one person's memory of an event can be influenced by that of another person, and that this post-event information can form a partly or completely new memory (e.g., Gabbert et al., 2004; Wright et al., 2000). The reason for this occurrence is that memory is very flexible and can be influenced by input occurring between encoding and retrieval (Wright & Loftus, 2008). The types of mechanisms that underlie memory blends are not clear. The trace-alteration hypothesis, for example, argues that the original memory is altered following the introduction of misinformation, whereas the source-monitoring view argues that participants fail to separate what they have experienced themselves and what has been said in the misinformation (Holliday et al., 1999). According to the source-monitoring view, people may remember a blend between an observation and a discussion about that observation, and the source of the final memory may be lost. Thus, people may remember something but may not remember that that memory originates partly from a post-event discussion about an event, not solely from the event itself. In terms of our current research participants may have remembered, for example, the colour of the person's scarf but, following discussion with a second person claiming that the person was wearing a scarf of a different colour, this information may have blended with the original information forming an intermediate response. This is in line with research (e.g., Azuma, Williams, & Davie, 2004; Belli, 1988; Loftus, 1977) showing that separate information sources activated in relative proximity to each other can induce memory blends. Alternatively, in our study the participants may have assumed that the colour of the scarf must have been a blend between the colour remembered by them and the colour remembered by their co-witness, or they may have purposefully given a compromise response in order to be more in line with their co-witness and in order to seem more accurate. Still, even if a memory report turns out to be a deliberate compromise response, this negatively affects eyewitness reliability.

In terms of future research, a focus should be put on establishing the nature of the processes underlying memory blends and whether they are mainly conscious or unconscious. It would also be important to try to estimate the frequency of memory blends in real eyewitnesses as well as establishing how the memory blend effect can be affected by individual differences and task demands.

Manuscript received 9 March 2007
Manuscript accepted 22 February 2008
First published online 10 April 2008

REFERENCES

Azuma, T., Williams, E. J., & Davie, J. E. (2004). Paws + cause = pause? Memory load and memory blends in homophone recognition. *Psychonomic Bulletin & Review*, *11*, 723–728.

Belli, R. F. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses? *Memory & Cognition*, *16*, 314–326.

Burt, C. D. B. (1999). Categorisation of action speed and estimated event duration. *Memory*, *7*, 345–355.

Burt, C. D. B., & Popple, J. S. (1996). Effects of implied action speed of estimation of event duration. *Applied Cognitive Psychology*, *10*, 53–63.

Doyle, J. M. (2005). *True witness: Cops, courts, science, and the battle against misidentification*. New York: Palgrave Macmillan.

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, *9*, 215–227.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps towards influence during a discussion. *Psychonomic Bulletin & Review*, *13*, 480–485.

Granhag, P., Ask, K., & Rebelius, A. (2005). *"I saw the man who killed Anna Lindh": A case study of eyewitness descriptions.* Paper presented at the European Law and Psychology Meeting, Vilnius, June.

Holliday, R. E., Douglas, K. M., & Hayes, B. K. (1999). Children's eyewitness suggestibility: Memory trace strength revisited. *Cognitive Development*, *14*, 443–462.

Lindsay, S. D. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 1077–1083.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560–572.

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition*, *5*, 696–699.

Loftus, E. F. (2004). Memories of things unseen. *Current Directions in Psychological Science*, *13*, 145–147.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*, 361–366.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.

Macdonald, P. D. M., & Green, P. E. J. (1988). *User's guide to program MIX: An interactive program for fitting mixtures of distributions. Release 2.3, January 1988*. Hamilton, Ontario: Ichthus Data Systems.

Memon, A., & Wright, D. B. (1999). The search for John Doe 2: Eyewitness testimony and the Oklahoma bombing. *The Psychologist*, *12*, 292–295.

Metcalfe, J. (1990). A composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General*, *119*, 145–160.

Mori, K. (2007). A revised method for projecting two different movies to two groups of viewers without their noticing the duality. *Behavior Research Methods*, *39*, 574–578.

Norman, D. A., Rumelhart, D. E., & the LNR Research Group (Eds.).(1975). *Explorations in cognition*. San Francisco: Freeman.

Paterson, H. M., & Kemp, R. I. (2006a). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime, & Law*, *12*, 181–191.

Paterson, H. M., & Kemp, R. I. (2006b). Comparing methods of encountering post-event information: The power of co-witness suggestion. *Applied Cognitive Psychology*, *20*, 1083–1099.

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. New York: Houghton Mifflin.

Schooler, J. W., & Tanaka, J. W. (1991). Composites, compromises, and CHARM: What is the evidence for blend memory representations? *Journal of Experimental Psychology: General*, *120*, 96–100.

Skagerberg, E. M., & Wright, D. B. (2008). Manipulating power can affect memory conformity. *Applied Cognitive Psychology*, *22*, 207–216.

Skagerberg, E. M., & Wright, D. B. (in press). The prevalence of co-witnesses and co-witness discussions in real eyewitnesses. *Psychology, Crime, & Law*.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA: Sage Publications.

Wright, D. B., & Loftus, E. F. (1998). How misinformation alters memories. *Journal of Experimental Child Psychology*, *71*, 155–164.

Wright, D. B., & Loftus, E. F. (2008). Eyewitness memory. In G. Cohen & M. A. Conway (Eds.), *Memory in the real world (3rd ed.)* (pp. 91–106). Hove, UK: Psychology Press.

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology*, *91*, 189–202.

Wright, D. B., & Skagerberg, E. M. (2008). Examining mixed unknown distributions (mud). *Behavior Research Methods*, *40*, 73–83.