

The response order effect: People believe the first person who remembers an event

Daniel B. Wright · Marianna E. Carlucci

Published online: 12 April 2011
© Psychonomic Society, Inc. 2011

Abstract When groups of people remember an event, the order in which they discuss their memories is important. In three experiments, a response order effect was shown in which participants believed the first speaker to be more accurate and more confident than a subsequent speaker. Further, participants were more likely to report as their own memory what the first speaker reported than what a subsequent speaker reported. The experiments showed that the response order effect was not due to intrinsic characteristics of what the first speaker said. Even when participants chose the response order themselves and the speakers' dialogue was counterbalanced, participants still believed that the first speaker was more accurate and confident than a subsequent speaker. Because in most situations the person who introduces a particular topic into a discussion is more accurate, people may assume that this is true, even when the response order is random.

Keywords Memory · Suggestibility · Memory conformity

Many events are witnessed by groups of people who later try to remember the event together. Examples include families remembering past holidays, friends remembering who bought the last round of drinks, students remembering a professor's lecture, and eyewitnesses to crimes remembering the culprit's appearance. When people remember in groups, one consequence is that what one person says can affect what other people believe and subsequently report. Over the past decade, several researchers (e.g., Af Hjelmsäter, Granhag,

Strömwall, & Memon, 2008; Axmacher, Gossen, Elger, & Fell, 2010; Bodner, Musch, & Azad, 2009; Brown & Schaefer, 2010; Candel, Memon, & Al-Harazi, 2007; Gabbert, Memon, & Allan, 2003; Hope, Ost, Gabbert, Healey, & Lenton, 2008; Mori & Mori, 2008; Paterson & Kemp, 2006; Reysen, 2005; Roediger, Meade, & Bergman, 2001; Schneider & Watkins, 1996; Skagerberg, 2007; Wright & Schwartz, 2010) have shown that what one person in the group reports can affect what other group members report (for reviews, see Blank, 2009; Wright, Memon, Skagerberg, & Gabbert, 2009). This is called *memory conformity* (Roediger, 2010; Wright, Self, & Justice, 2000).

Memory conformity research is similar to the classic postevent information (PEI) research (Loftus, 2005). The main difference is that in memory conformity research, the PEI is introduced by another person who has seen the same event (sometimes confederates are used, and sometimes slightly different events are shown to the different participants). This difference introduces some important variables that are not part of the classic PEI studies. Wright, London, and Waechter (2010) described how responses in these memory conformity situations were affected by both normative influences (i.e., the importance of giving the same response as others) and informational influences (i.e., beliefs about the accuracy of your own and other people's memories). The focus of this article is on informational influences, and in particular on how people weight the reliability of other people's memory reports, depending on whether the person responds first or second.

Gabbert, Memon, and Wright (2006) examined response order and memory conformity. Pairs of participants were shown complex scenes in which some of the details were different. The pairs discussed the scenes and then were tested individually. Gabbert et al. (2006) found that participants who first brought up the critical detail during

D. B. Wright (✉) · M. E. Carlucci
Psychology Department, Florida International University,
11200 S. W. 8th Street,
Miami, FL 33199, USA
e-mail: dwright@fiu.edu

the discussion tended to influence the other person in the pair more than they were influenced themselves. Allowing a natural dialog is important for ecological validity, but such a procedure makes it difficult to understand why this effect occurs (Lindsay, 2007).

In Gabbert et al. (2006), the observed association between the response order and memory conformity could have been due to one of three reasons:

1. Some aspect of the dialogue is different when people speak first than when people speak second. It may be that the person who introduces the topic uses more confident language.
2. People may believe that if a person *chooses* to speak first, then that person is likely to be more confident and more accurate than a person who did not choose to speak first. Given that in natural dialogue the tendency is for the more confident person to initiate discussion, this would be an accurate belief.
3. People may assume that those who speak first, regardless of the reason for the response order, will be more confident and more accurate, because this heuristic works appropriately in most everyday situations in which members of the group determine the response order.

These three reasons were examined here in three experiments. In Experiment 1, we tested whether the response order effect remains even when the actual memory reports are identical. In Experiment 2, we tested whether the effect remains when participants could see that the response order was not initiated by the speakers. Experiments 1 and 2 used within-subjects designs and had the participants also view the stimuli. In Experiment 3, the response order effect was tested with a between-subjects design with a large sample in which participants were randomly allocated to one of three conditions. Some were told that the response order was decided by the speakers, some that the response order was random, and some of the participants chose the response order themselves.

Experiments 1 and 2

In the first two experiments we used similar methods, so here they are described together. Participants were presented with a series of pictures and told that this was part of a study in which other people saw these same pictures. Participants were told that they were matched with a pair of participants who took an old/new recognition test together. For each trial at testing, participants watched a videotape of the pair responding and then responded with their own answer. The most informative trials for evaluating the response order effect occurred when the first two people disagreed. The difference between the two experiments was

whether the response order was initiated by one of the confederates by ringing a bell (Exp. 1) or was random, initiated by a light that went off (Exp. 2).

Method

Participants

A total of 200 participants (81 in Exp. 1, 119 in Exp. 2) from Florida International University's psychology participant pool took part in exchange for class credit. Their mean age was 20 years old, most were female (61%), and most were Hispanic (71%).

Materials

Stimuli The stimuli were 100 complex black and white drawings taken from a large set (Cirker, 2006; Dover Electronic Clip Art, 2007). Examples are shown in Fig. 1.

Videos The videos showed two males answering "new" or "old" for 100 trials in the same room in which the participants were later tested. In Experiment 1, one of the males rang a "call bell," like those at hotel front desks, and then answered. In Experiment 2, a green bulb lit up to alert one of the males to answer. The videos were recorded such that both confederates were visible.

Procedure

Participants sat in front of a computer and were shown 50 pictures for 1 s each using the SuperLab software. Next, participants were given an answer sheet and told that they

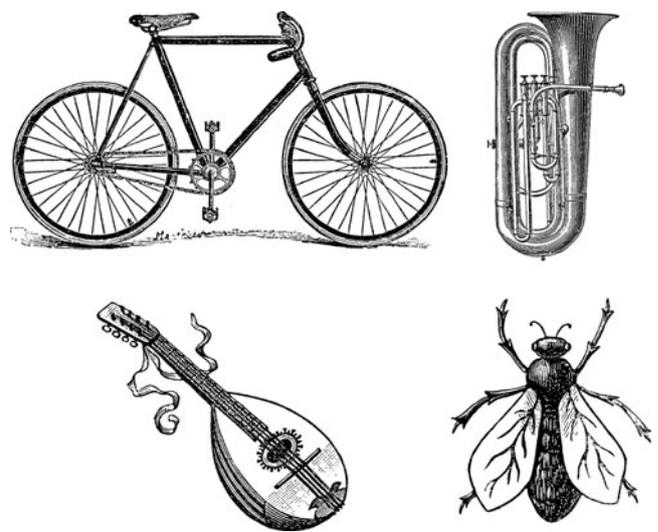


Fig. 1 Examples of the stimuli used in Experiments 1 and 2

would be viewing 100 pictures, and that some would be lures (“never seen before”) and some would be targets (“previously seen”). For each of the 100 pictures, participants saw a video of two other “participants” (confederates) responding to that picture. The participants in each video either rang a bell to answer (Exp. 1) or had a green bulb light up in front of them when it was their turn to answer (Exp. 2). The person who rang the bell (Exp. 1) or whose bulb lit up (Exp. 2) answered first. The two people gave the same correct answer 40% of the time, the same incorrect answer 20% of the time, and disagreed 40% of the time (half of these times, the first speaker was correct). These percentages were the same for old items (targets) and new items (lures). Each confederate responded the same number of times in each condition. The videos were counterbalanced.

After viewing a video, the participants were asked to make a decision about the picture (“new” or “old”). They were also asked to make confidence judgments about the two confederates’ answers using a 1 (*not confident*) to 10 (*very confident*) scale. Each participant went through all 100 trials in succession. The entire procedure took approximately 45 min.

Results

In the standard old/new recognition procedure, memory is shown if people are more likely to say “old” to targets (i.e., hits) than to say “old” to lures (i.e., false alarms). Statistically, these data can be analyzed in several ways (Wright, Horry, & Skagerberg, 2009). Two common statistics for this procedure are the *log-odds ratio* [i.e., $\text{logit}(\text{hit rate}) - \text{logit}(\text{false alarm rate})$] and d' [i.e., $\text{probit}(\text{hit rate}) - \text{probit}(\text{false alarm rate})$]. The logit and the probit are different ways of transforming the hit rates and false alarm rates to make them more amenable to statistical analysis, because they are not bounded by 0 and 1 as proportions are. These transformations are widely used in signal detection theory, and they usually produce similar results. Each is also related to a type of generalized linear model. The log-odds ratio can be calculated using a logistic regression, while d' is estimated by a probit regression. When the analysis is framed as a regression model, other predictor variables can be included. Memory conformity can be measured by including a variable for the confederate’s response and testing how much participants’ responses are affected by the confederates’ responses. Using terminology from signal detection theory, the confederate’s response can be thought of as moving the threshold for responding “old” or “new” (Wright, Gabbert, Memon, & London, 2008). The interaction between condition and the confederate’s response measures whether condition moderates the memory conformity effect.

Because each person makes multiple judgments, estimating memory and memory conformity is more complex in a within-subjects design than when a between-subjects design is used. Following Baayen, Davidson, and Bates (2008); (see also Goldstein, 2003; Wright & London, 2009), we used multilevel logistic/probit regressions. The R package lme4 (Bates & Maechler, 2010) was used for all analyses, with random intercepts for both items and participants included. Statistical tests were done using logistic regressions, but both the log-odds ratio and d' are presented below for the effect sizes.

Table 1 shows the proportion of “old” responses to targets and lures, dependent on what the confederates said for both experiments. The differences between the proportions in the “Targets” and the “Lures” columns (40%–50% shifts) show that participants had accurate memories for the pictures. The differences between the columns for what the second speaker said show small effects for what the second speaker said (–2% to 8% shifts). The differences between the rows for what the first speaker said show larger effects (7%–17% shifts).

For Experiment 1, when the confederates agreed, participants’ responses were predicted from both whether the item was a target or a lure and whether the confederates said “old” or “new.” Both of these effects were statistically significant: for memory, $\chi^2(1) = 85.78, p < .001$, and for memory conformity, $\chi^2(1) = 29.94, p < .001$. The first units for the effect sizes of logistic regressions are in log-odds ratios (lnOR), and the second in d' . The effect for memory was $\text{lnOR} = 2.32 (SE = 0.17)$ and $d' = 1.39 (SE = 0.10)$. The effect for memory conformity was $\text{lnOR} = 1.06 (SE = 0.17)$ and $d' = 0.62 (SE = 0.10)$. The corresponding analyses for Experiment 2 were $\chi^2(1) = 87.29, p < .001$, for memory and $\chi^2(1) = 34.10, p < .001$, for memory conformity. The effect

Table 1 Proportions of “old” responses as a function of whether the item was old or new and what the two confederates said in Experiment 1 (bell cue) and in Experiment 2 (light cue)

	Lures (False alarm rate)		Targets (Hit rate)	
	Second says “new”	Second says “old”	Second says “new”	Second says “old”
Experiment 1				
First confederate responds “new”	.14	.19	.63	.68
First says “old”	.28	.36	.74	.78
Experiment 2				
First confederate responds “new”	.16	.19	.67	.65
First says “old”	.30	.36	.74	.78

for memory was $\ln\text{OR} = 2.21$ ($SE = 0.15$) and $d' = 1.33$ ($SE = 0.09$). The effect for memory conformity was $\ln\text{OR} = 1.03$ ($SE = 0.15$) and $d' = 0.59$ ($SE = 0.09$).

The focus of this article is when the two confederates disagreed. From Table 1, in Experiment 1 with lures, participants said “old” 28% of the time when the first speaker had said “old” but the second speaker had said “new.” When the response order was reversed, only 19% of participant responses were “old.” For targets, if the first speaker had said “old” and the second said “new,” the participants said “old” 74% of the time, but they said “old” only 68% of the time when the response order was reversed. Overall, when the speakers disagreed the participant gave the same answer as the first speaker 54% of the time and the same answer as the second speaker 46% of the time. In Experiment 2, for lures, if the first speaker had said “old” but the second speaker said “new,” the participant said “old” 30% of the time. When this response order was reversed, only 19% of participant responses were “old.” For targets, if the first speaker had said “old” and the second speaker said “new,” the participants said “old” 74% of the time. When the response order was reversed, participants said “old” only 65% of the time. Overall, when the speakers disagreed, the participant gave the same response as the first speaker 55% of the time, and the same answer as the second speaker 45% of the time.

Inferential statistics for trials on which the speakers disagreed were analyzed using multilevel logistic and probit regressions to predict the participants’ responses from whether the item was a target or a lure and from what the first speaker said. If the parameter associated with the first speaker’s response was positive, that meant that the participant tended to give the same response as the first speaker. If the parameter was negative, that meant that the participant tended to give the same response as the second speaker. For Experiment 1, both memory and the tendency to agree with the first speaker were statistically significant: for memory, $\chi^2(1) = 66.84$, $p < .001$, and for agreeing with the first speaker, $\chi^2(1) = 5.62$, $p = .02$. For the memory effect, $\ln\text{OR} = 2.37$ ($SE = 0.18$) and $d' = 1.42$ ($SE = 0.11$), so this effect was similar to the estimate for memory found when the speakers agreed. The size of the response order effect was $\ln\text{OR} = 0.45$ ($SE = 0.18$) and $d' = 0.26$ ($SE = 0.11$), which was positive and therefore confirmed the observations from Table 1 that participants tended to go with the first speaker. Experiment 2 produced similar statistics: The effect for memory was $\ln\text{OR} = 2.18$ ($SE = 0.17$), $d' = 1.31$ ($SE = 0.10$), $\chi^2(1) = 63.60$, $p < .001$, and the response order effect was $\ln\text{OR} = 0.57$ ($SE = 0.17$), $d' = 0.34$ ($SE = 0.10$), $\chi^2(1) = 9.54$, $p = .002$.

Participants provided confidence ratings for their own responses and those of the two people in the video. In Experiment 1, participants gave high ratings for their own confidence (mean = 8.27 out of 10), high ratings for each

confederate when they agreed with them (8.18), and lower ratings for a confederate with whom they disagreed (5.68). We focused on the participants’ own confidence when the confederates disagreed. When participants agreed with the first responder, the mean confidence was 8.28, versus 8.06 when they agreed with the second responder. Because of the ceiling effect, a generalized multilevel model for proportions was used, taking into account that the confidence score was out of 10. We allowed a random intercept for participants but not a random term for pictures, since the latter estimate was very small (variance = .001). The analysis showed that this difference was statistically significant, $\chi^2(1) = 4.28$, $p = .04$.

The results for the confidence ratings were similar for Experiment 2. Overall, participants gave high ratings to their own responses (mean = 8.23) and to confederates whom they agreed with (8.07), but lower ratings to confederates whom they disagreed with (5.45). When confederates disagreed, if the participant went with the first responder’s answer, the participant’s confidence was higher than when he or she agreed with the second responder (8.36 vs. 8.08). Using the same methods as with Experiment 1, this difference was statistically significant, $\chi^2(1) = 6.10$, $p = .01$.

Discussion

Overall, the effects were similar for the two experiments. Memory was nearly equivalent across experiments and across whether the speakers agreed or disagreed. The memory conformity effects were smaller than the effects for memory. The effects of most interest were that participants conformed more with the first than with the second speaker. In Gabbert et al. (2006), the response order effect could have been caused solely by something within the first speaker’s dialogue that was more convincing. By showing each speaker in each order in our new study, this hypothesis could be rejected. The effects in Gabbert et al. (2006) could have been caused by participants believing that people who choose to speak first do so because they are more confident, and therefore more accurate, than people who do not choose to speak first. Experiment 2 addressed this hypothesis. The response order effect remained, even when participants could see that the response order was not initiated by the responders.

The participants’ task in Experiments 1 and 2, making 100 judgments, was complex. It may be that some of the participants in Experiment 2 did not take into account that the response order was random because of having to respond to so many trials. It may be that if participants engaged in only one trial that the manipulation would be more salient. Therefore, in the Experiment 3 participants

made only a single judgment. Because of having only a single judgment, a larger sample is necessary.

In addition, in order to make the participant believe that the two speakers did not choose their own response order, it might be necessary to ask the participants themselves to choose the order. Therefore, in Experiment 3 a random sample of the participants were told to choose the response order of the two speakers.

Finally, in Experiments 1 and 2 participants would have memories for some of the stimuli. One question is whether the response order effect would generalize to situations in which the person was not shown the event. In Experiment 3, participants were not shown the original event, but were simply asked to judge the confidence and accuracy of the two speakers.

Experiment 3

Method

Sample

A large sample was needed as a result of using a between-subjects design in which each participant responded to only one trial. The online survey company, Qualtrics (www.qualtrics.com), was employed to recruit participants from their existing panel of over four million people. E-mails were sent to a subset of people who could volunteer to participate. Qualtrics uses different filters—for example, removing people who do not spend adequate time on Web pages and who do not tick boxes as instructed. Here, participants were also excluded if they could not see or hear the videos on their computer system. There were 2,125 participants (70% female, 82% White, and 50% between 25 and 54 years old). Qualtrics paid each participant a small fee, depending on how many surveys they completed.

Materials

Four videos were made. Two female confederates each memorized two scripts about a restaurant scene. The scripts were the same except for a few details, including the type of restaurant, customers' drink orders, and who paid the bill. The confederates were videotaped giving their "recollection" of the scene for both scripts. The participants saw two videos, one with each confederate and one with each script (counter-balanced). The videos were in .flv (Flash video) format, since this format can be viewed by more operating systems than other formats (this is the format used by YouTube). These videos were created using the Prism video converter software (www.nchsoftware.com/prism).

Procedure

The experiment was conducted online using the Qualtrics software. Participants were told that they would see two videos of people describing a restaurant scene. Participants were randomly allocated to one of three conditions. In the first condition, they were told that the two speakers decided which person would go first (actor decides). In the second condition, they were told that the computer randomly chose the response order (computer decides). In the third condition, the participants chose the response order themselves, with either Susan or Amanda first (participant decides). The names were presented in random order, and the participant did not see a photograph or any other information about either speaker. In the third condition, the participant's choice did not determine which video was actually shown first. As a manipulation check, participants were then asked how the response order was determined and were excluded if they incorrectly answered this.

Participants were told to make sure that the computer speakers were turned on and to press "play." Which of the videos was shown first was random. The second video was of the speaker not in the first video using the other script. Participants watched the videos and were then asked whether they could see and hear the videos (and were excluded otherwise).

Next, participants rated how accurate they thought the first and second speakers were and how confident they felt the first and second speakers were. All ratings used a 1–10 scale. Responses to these four questions were the dependent variables.

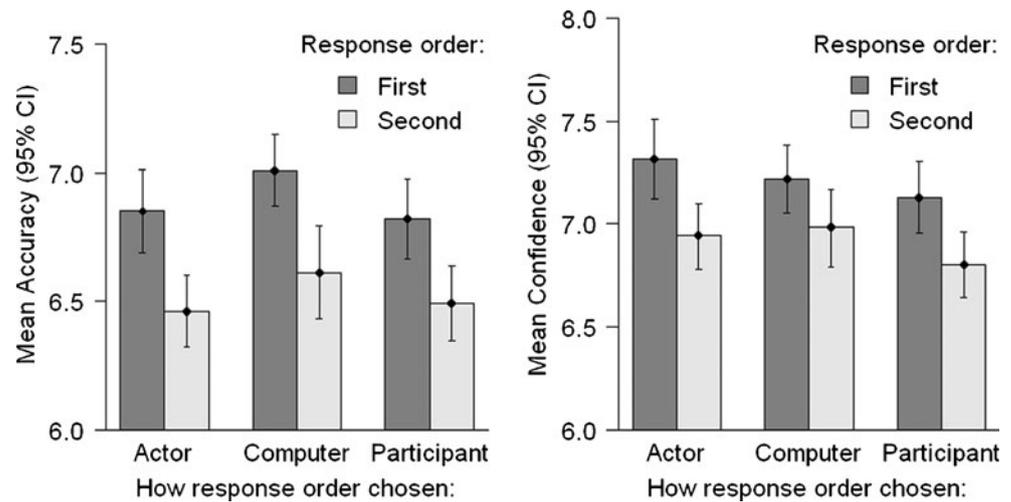
Results and discussion

Eighty-two percent of participants (1,732 of 2,125) correctly stated which condition they were in and that they were able to see and hear the videos. The remaining 18% were excluded.

Figure 2 shows the mean ratings for accuracy and for confidence for the first and the second speaker by participant condition. Table 2 provides the ANOVA statistics. Main effects were found for response order, with the first speaker being judged as more accurate and more confident. There were no significant main effects for condition and no significant interactions. Differences between ratings for the first and second speakers for each measure were subjected to a paired *t* test, and each was statistically significant ($p < .05$), with higher ratings for the first speaker for all six comparisons.

Variables were created for the difference between the rating for the first speaker minus the rating for the second speaker, for both accuracy and confidence. The prediction that people who felt that one speaker was more confident

Fig. 2 Mean ratings of accuracy (left panel) and confidence (right panel), with their 95% between-subjects confidence intervals, for each condition



also felt that that speaker was more accurate were confirmed, $r = .69$ ($n = 1,732$), 95% CI from .66 to .72.

Even when participants chose the order of the speakers, they thought that the first speaker was more accurate and more confident. The interactions were all nonsignificant. Because participants only made a single judgment, the statistical tests were not as powerful as if they had made many judgments, as in Experiments 1 and 2, but given the sample size we are confident that if any differences existed among the conditions for the response order effect, they were small.

General discussion

When groups of people witness the same event, they often try to remember the event together. People listening to such a discussion will be affected by it, which is called *memory conformity* (Wright et al., 2009). If all the speakers agree, the influence will be large. If the speakers are all accurate, this will increase the likelihood that subsequent responders will be accurate, and if they are all inaccurate, this will decrease the likelihood that subsequent responders will be accurate. The focus of this article was what happens when

pairs of speakers disagree. Gabbert et al. (2006) reported data showing an association between response order and influence. They found that when pairs of people responded, the first person who brought an item up usually influenced the other person. Lindsay (2007) described some of the methodological limitations in trying to distinguish the causal factors from Gabbert et al. (2006) study. The aims of the present research were to tease apart some of the confounding variables in their design, to show whether the response order effect remains even when people believe that the response order is random, and to generalize the effect to situations not involving the participant's memory.

The findings from the three experiments show that the response order effect does remain even when people view the same videos and believe either that the response order is random or that they determined it themselves. While in Experiment 1 and the "Actor decides" condition of Experiment 3 participants might believe that the first speaker was more accurate and confident than the second speaker because the first speaker *chose* to respond first, the size of the response order effects were similar in Experiments 1 and 2 and did not vary significantly among the conditions of Experiment 3. When participants believed either that the computer randomly decided the response order or that the participants decided the response order themselves, participants should logically not believe a priori that the first speaker would be more confident and more accurate, but they did.

The conclusion from these experiments is that participants believe that when two people each recount some event that the first speaker is more accurate than the second speaker. There are different explanations for this effect. One is that people are accustomed to everyday situations in which the first speaker tends to be more accurate and more confident, and that they have developed a heuristic to weight the first speaker as more reliable. When people freely report their memories, they are usually very

Table 2 ANOVA statistics for Experiment 3

Accuracy			
First versus second	$F(1, 1729) = 115.33$	$p < .001$	$\eta_p^2 = .027$
Condition	$F(2, 1729) = 2.04$	$p = .13$	$\eta_p^2 = .002$
Interaction	$F(2, 1729) = 0.18$	$p = .97$	$\eta_p^2 = .000$
Confidence			
First versus second	$F(1, 1729) = 22.50$	$p < .001$	$\eta_p^2 = .013$
Condition	$F(2, 1729) = 1.83$	$p = .16$	$\eta_p^2 = .002$
Interaction	$F(2, 1729) = 0.35$	$p = .71$	$\eta_p^2 = .000$

confident and accurate (Koriat & Goldsmith, 1996), and therefore in most everyday situations this heuristic will be useful. When somebody responds first in a dialogue, even an artificial, laboratory dialogue in which the response order is random, observers do not appear to take into account why one person responds first. People use the usually accurate heuristic that the person who responds first is more likely to be correct, and therefore weight that person's response as more valid than subsequent people's responses. There may even be situations in which participants should weight the second speaker as more reliable. Suppose that the first speaker says that a person was wearing a cowboy hat. If the second speaker said: "No, it was a baseball hat," the fact that the second speaker is disagreeing suggests that the second speaker may be very confident.

There are many situations in which people are presented with contradictory information from two sources. Future research will be necessary to explore whether the response order effect will generalize beyond people recounting a memory. For example, in a political debate, two politicians often describe a situation in very different ways. Another important avenue for exploration is what occurs if the time between the speakers is lengthened. Understanding more about how the response order of speakers affects how their statements are encoded and interpreted is an important topic for understanding cognition within a social context.

Author Note Thanks to Erika Fountain, Holly George-John, Michael Leduc, Jennifer Mejias, and Crystal Mueller for helping with data collection. Thanks to Osvaldo de la Cruz and Jan Garcia for acting in Experiments 1 and 2, and to Joemy Familia and Alyssa Ruiz for acting in Experiment 3. Experiment 3 was funded by support from Florida International University's College of Arts and Science.

References

- Af Hjelsäter, E. R., Granhag, P. A., Strömwall, L. A., & Memon, A. (2008). The effects of social influence on children's memory reports: The omission and commission error asymmetry. *Scandinavian Journal of Psychology*, *49*, 507–513. doi:10.1111/j.1467-9450.2008.00679.x
- Axmacher, N., Gossen, A., Elger, C. E., & Fell, J. (2010). Graded effects of social conformity on recognition memory. *PLoS ONE*, *5*, e9270. doi:10.1371/journal.pone.0009270
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005
- Bates, D., & Maechler, M. (2010). lme4: Linear mixed-effects models using S4 classes (R package version 0.999375-33). Available at <http://CRAN.R-project.org/package=lme4>
- Blank, H. (2009). Remembering: A theoretical interface between memory and social psychology. *Social Psychology*, *40*, 164–175. doi:10.1027/1864-9335.40.3.164
- Bodner, G. E., Musch, E., & Azad, T. (2009). Reevaluating the potency of the memory conformity effect. *Memory & Cognition*, *37*, 1069–1076. doi:10.3758/MC.37.8.1069
- Brown, C., & Schaefer, A. (2010). The effects of conformity on recognition judgements for emotional stimuli. *Acta Psychologica*, *133*, 38–44. doi:10.1016/j.actpsy.2009.08.004
- Candel, I., Memon, A., & Al-Harazi, F. (2007). Peer discussion affects children's memory reports. *Applied Cognitive Psychology*, *21*, 1191–1199. doi:10.1002/acp.1355
- Cirker, B. (Ed.). (2006). *1268 old-time cuts and ornaments*. Mineola: Dover.
- Dover Electronic Clip Art. (2007). *1565 spot illustrations and motifs*. Mineola: Dover.
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology*, *17*, 533–543.
- Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*, 480–485.
- Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Edward Arnold.
- Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). "With a little help from my friends...": The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica*, *127*, 476–484. doi:10.1016/j.actpsy.2007.08.010
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490–517.
- Lindsay, D. S. (2007). Order effects in collaborative memory contamination? Comment on Gabbert, Memon, and Wright (2006). *Psychonomic Bulletin & Review*, *14*, 1010.
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*, 361–366. doi:10.1101/lm.94705
- Mori, K., & Mori, H. (2008). Conformity among cowitnesses sharing same or different information about an event in experimental collaborative eyewitness testimony. *Perceptual and Motor Skills*, *106*, 275–290.
- Paterson, H. M., & Kemp, R. I. (2006). Comparing methods of encountering post-event information: The power of co-witness suggestion. *Applied Cognitive Psychology*, *20*, 1083–1099. doi:10.1002/acp.1261
- Reysen, M. B. (2005). The effects of conformity on recognition judgements. *Memory*, *13*, 87–94. doi:10.1080/09658210344000602
- Roediger, H. L., III. (2010). Reflections on intersections between cognitive and social psychology: A personal exploration. *European Journal of Social Psychology*, *40*, 189–205. doi:10.1002/ejsp.736
- Roediger, H. L., III, Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, *8*, 365–371.
- Schneider, D. M., & Watkins, M. J. (1996). Response conformity in recognition testing. *Psychonomic Bulletin & Review*, *3*, 481–485.
- Skagerberg, E. M. (2007). Co-witness feedback in lineups. *Applied Cognitive Psychology*, *21*, 489–497. doi:10.1002/acp.1285
- Wright, D. B., Gabbert, F., Memon, A., & London, K. (2008). Changing the criterion for memory conformity in free recall and recognition. *Memory*, *16*, 137–148. doi:10.1080/09658210701836174
- Wright, D. B., Horry, R., & Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, *41*, 257–267. doi:10.3758.BRM.41.2.257
- Wright, D. B., & London, K. (2009). Multilevel modelling: Beyond the basic applications. *The British Journal of Mathematical and Statistical Psychology*, *62*, 439–456. doi:10.1348/000711008X327632

- Wright, D. B., London, K., & Waechter, M. (2010). Social anxiety moderates memory conformity in adolescents. *Applied Cognitive Psychology, 24*, 1034–1045. doi:10.1002/acp.1604
- Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science, 18*, 174–178.
- Wright, D. B., & Schwartz, S. L. (2010). Conformity effects in memory for actions. *Memory & Cognition, 38*, 1077–1086. doi:10.3758/MC.38.8.1077
- 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.