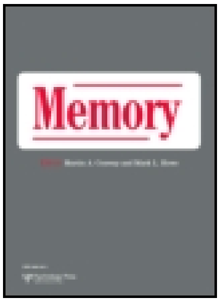

Memory
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pmem20

Judging the reality of others' memories

Arlo Clark-Foos[a], Gene Brewer[b] & Richard L. Marsh[c]

[a] Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn, MI, USA

[b] Department of Psychology, Arizona State University, Tempe, AZ, USA

[c] Department of Psychology, University of Georgia, Athens, GA, USA

Published online: 10 Mar 2014.

Memory, 2014
http://dx.doi.org/10.1080/09658211.2014.893364

Judging the reality of others' memories

Arlo Clark-Foos[1], Gene Brewer[2], and Richard L. Marsh[3†]

[1]Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn, MI, USA
[2]Department of Psychology, Arizona State University, Tempe, AZ, USA
[3]Department of Psychology, University of Georgia, Athens, GA, USA

(*Received 30 April 2013; accepted 9 January 2014*)

Interpersonal reality monitoring (IRM) refers to our ability to evaluate whether other people's memories reflect real or imagined events. The current work examined IRM and whether or not it can be affected by training and feedback. We found that people are better than chance and that the ability to accurately make this judgement can be improved or reduced with appropriate and inappropriate training, respectively. Understanding IRM has implications for applied psychologists interested in how people evaluate others' descriptions of past events (e.g., eyewitness testimony).

Keywords: Interpersonal reality monitoring; Memory.

The proposition that an individual can accurately judge whether another individual experienced or imagined an event based solely on a description of the memory for that event is tantalising and important to consider for a variety of applied and theoretical reasons. For example, in applying psychological principles to eyewitness testimony, it is critical to account for the inferential processes that drive jurors' assessments of descriptions of events in the past (Bell & Loftus, 1989). Theoretically speaking, distinguishing the processes that are necessary for accurate inter-individual memory monitoring [or interpersonal reality monitoring (IRM; Johnson, Bush, & Mitchell, 1998)] is just as essential. To preface, we are attempting to address these issues using basic experimental research motivated by the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993). In the current study, we ask whether participants can determine if another person's written descriptions of their memories represent real or imagined events. Moreover, we question whether people can use training and feedback to improve the accuracy of their judgments of the source of others' memories.

Reality monitoring refers to the collection of processes used to determine whether our own memories are formed by actual experience or mere imagination (Johnson & Raye, 1981). J. S. Mill wrote that "the only difficulty about Memory … is the difference between it and imagination" (1869/1967, p. 339). He was lamenting the fact that memories for real and imagined events are very similar. While we are generally successful at monitoring the output from our own memory system, it is currently unclear how accurate we are at monitoring someone else's memory (for notable exception, see Johnson & Suengas, 1989).

Address correspondence to: Arlo Clark-Foos, Department of Behavioral Sciences, University of Michigan, Dearborn, MI 48128, USA. E-mail: acfoos@umich.edu

We thank Alexander Diaz, Melanie Woodward, Katherine Pfannes and Michelle Evans for their assistance in collecting the data. We also thank Brian Cutler for helpful comments on an earlier version of this manuscript.

†Deceased 8 June 2010.

Reality monitoring is a special case of source monitoring, which is a set of processes that an individual uses to determine the origin of memories. The source-monitoring framework (Johnson et al., 1993) proposes that we assess both the quality and quantity of information in a target "memory" to determine the source of that memory. A memory is thought of as a set of features (Bower, 1967), each representing different attributes of the original encoding experience, then judgments of origin can be based on qualitative or quantitative differences in retrieved features. However, this process is subject to error because activated features may arise not only from the target event itself, but also from other sources (e.g., other, similar events, imaginations, etc.; e.g., Henkel, Franklin, & Johnson, 2000; Lyle & Johnson, 2006). In many reality-monitoring situations, one can identify the origin of a memory by retrieving a single qualitative feature that is indicative of source (a point we will return to later). In the absence of a discriminative feature, however, it may be necessary to use a more quantitative weighting of features to determine which source is more likely (Hicks, Marsh, & Ritschel, 2002).

When making reality-monitoring decisions, one rarely has access to a discriminative feature that accurately specifies the source of the memory. Instead, some quantitative assessment of retrieved features is made to determine which of several sources has more evidence. The relative amount of retrieved evidence (i.e., memory attributes) will be compared with one's "metamemory" assumptions about what constitutes real and imagined memories (Foley, Johnson, & Raye, 1983; Johnson & Raye, 1981). We are not, however, saying that qualitative information cannot be used to make reality-monitoring judgments. It is probable that most reality-monitoring decisions are made using a combination of discriminative (source specific) features and/or weighting of multiple features that are probabilistically associated with different sources (Hashtroudi, Johnson, & Chrosniak, 1990). Such features include: spatial and temporal context, sensory details, semantic details, emotion and cognitive operations (Johnson et al., 1993). On average, memories for imagined events have more cognitive operations associated with generating the representation, and memories for real events have more sensory details (Finke, Johnson, & Shyi, 1988; Hashtroudi et al., 1990; Johnson et al., 1998; Schooler, Gerhard, & Loftus, 1986; Suengas & Johnson, 1988). Occasionally, our memories have discriminative features that indicate the likely sources with a high degree of certainty. In the absence of a discriminative feature, one can apply decision criteria to the dimensions of retrieved attributes to make attributions about the origin of a memory that should be better than chance (how much better than chance, of course, depends on the overlap in the distributions of features from various sources).

Just as source-monitoring failures result in making inappropriate attributions about the source of a memory, reality-monitoring failures result in making inappropriate attributions about the physical occurrence of a memory. There are certainly cases where one purposefully lies about the occurrence of an event (e.g., false eyewitness testimony). In fact, a wealth of literature exists that examines the influence of deception on judging the accuracy of another person's memory (see Masip, Sporer, Garrido, & Herrero, 2005 for a review of studies using the RM approach to detecting deception). How often, however, does it occur that someone believes an event to have existed when it actually has not, and vice versa? For example, we know that people can report memories for events from childhood that they never actually experienced but merely imagined (e.g., Hyman & Loftus, 2002). At the same time, participants will not form a false memory for an unlikely event, even if they are capable of vividly imagining it (Mazzoni, Scoboria, & Harvey, 2010). When these individuals inevitably report their memories to others, are we able to accurately reality monitor their output (for a related discussion, see Ost, Vrij, Costall, & Bull, 2002)? The particular attributions one makes about the source of another person's memory should affect the way that person is treated by others (Johnson, 1988), which is why our study represents a particularly interesting case of the reality-monitoring discrimination process. That is, we aim to test peoples' abilities in the scenario of evaluating a written description of a memory (cf. Barnier, Sharman, McKay, & Sporer, 2005; Keogh & Markham, 1998). In this study, we chose to remove the element of deception as it has been shown to affect how we make Interpersonal Reality Monitoring (IRM) judgments (Barnier et al., 2005; Vrij, 2000) and we are interested in IRM judgments of truthful accounts. Because IRM is affected by suspected deception and because we were interested in cases where the honesty of the rememberer was not in question (as is the case in reality-monitoring judgments), we told participants that all memories were reported honestly, without any intended deceit (cf. Johnson et al, 1998). That is, we were interested in how people judge the origin of another's memory (IRM) in the absence of any suspected deception.

EXPERIMENT 1

To the best of our knowledge, only a few other studies have looked directly at reality monitoring other peoples' memories for actually experienced, not rehearsed, events (Johnson & Suengas, 1989; Keogh & Markham, 1998; Schooler, Clark, & Loftus, 1988, Schooler et al., 1986; Sporer & Sharman, 2006). Participants in Johnson and Suengas's (1989) study were asked to rehearse either the perceptual or apperceptive (i.e., conceptual) features of ordinary events that they either imagined or actually witnessed in the lab (e.g., wrapping a box). That is, participants were asked to think about either perceptual (e.g., the colour of the wrapping paper) or conceptual (e.g., emotions experienced while witnessing) features of those situations. Following rehearsal, all participants were asked to write descriptions of what they could remember from the original event. Resulting memory for both real events receiving apperceptive rehearsal and imagined events receiving perceptual rehearsal was no longer able to be accurately discriminated. These data strongly suggest that memories for real and imagined events should differ qualitatively in the type of information that is recalled. Although memories for real and imagined events should differ in their references to sensory details or cognitive operations, these differences did not bear out in transcriptions of the rehearsed memories (Johnson & Suengas, 1989). On the one hand, people may embellish oral descriptions of memories, resulting in greater similarity in the oral descriptions of real and imagined events (c.f. Johnson & Suengas, 1989). On the other hand, writing about an imagined event may make the memory seem more concrete, increasing its similarity to real events (Johnson, 1988). It is important, therefore, to validate the current paradigm using written descriptions of other people's *unrehearsed* memories. Experiment 1 had two purposes: (1) collect written descriptions of memories for objects that were either real or imagined and (2) obtain an estimate of baseline IRM performance with these materials (i.e., without training or feedback).

Methods

Participants

Thirty participants from the University of Georgia participated for partial fulfilment of a course research requirement.

Materials

In order to generate descriptions of real and imagined objects, 20 different participants from the same subject pool were asked to view a slideshow containing five photographs of everyday objects (*boat, cabin, microwave, chair, hamburger, truck, cow, cake, penguin* and *mushroom*) and five words naming different everyday objects. Objects serving as photos versus words were counterbalanced so that each object appeared equal number of times as either a photo or a word. Photographs were chosen from a variety of online sources (e.g., Google). Each slide was displayed for two minutes and participants were asked to either study the photograph or form a mental image of the object named by the word. They were also told that they would later be asked to recall as much as they could about either the real or imagined image. A five-minute math distractor followed the slideshow, after which all participants were given a booklet containing 10 blank pages, with the name of one of the objects written at the top of each page (for a similar procedure, see Hashtroudi et al., 1990). Participants were given as much time as they needed to describe their experience of each object.[1]

Two participants' descriptions (20 total descriptions) were excluded because they could not recall any details about at least one of the items. All other descriptions (180) were cleaned to remove obvious references to the source of the memory (Johnson & Suengas, 1989). That is, we modified language that specifically named the origin of the memory (i.e., "I imagined…"). This process resulted in approximately 1% of words being changed or eliminated, with no differences in the number of words changed between real and imagined sources. Previous research (e.g., Johnson & Raye, 1981) has suggested that there are differences in the characteristics encoded with memories of real and imagined events. All 180 descriptions were examined for references to overall clarity,

[1] Pictures and participant descriptions are obtainable from the following website: http://www-personal.umd.umich.edu/~acfoos/resources/smdesc_stim.xlsx.

sensory details, spatial information, references to time, cognitive operations and emotion, using a procedure based on the Self-Ratings of Memory Characteristics Questionnaire (SMCQ; Sporer & Kuepper, 2004). The SMCQ includes over 40 dimensions to assess various aspects of a specific memory. Although the scales vary slightly from dimension to dimension (e.g., none–many, low–high, not precise–precise), all use a 0–2 point rating. Two raters, blind to description source, were engaged in this analysis and proportion of initial agreement on all factors exceeded 80%, with disagreements being resolved through discussion. No adjustment to ratings was made for overall length of the description. Sporer (1997, 2004) asked trained experts to rate transcriptions of real and fabricated events using a similar approach and found that the resulting ratings could be used to accurately (69%) classify the transcriptions. Analysis of differences in our memories for real and imagined items revealed only four significant differences, with more references to seasons (time of year) and doubts about accuracy in memories from real sources and more references to sounds and locations of objects found in memories from imagined sources, all $ps < .01$ (see Table 1 for means of all dimensions for real and imagined sources). More frequent remembering of seasons in real sources may be the result of some of our images having implied information about the season of the year (e.g., lush foliage surrounding a boat probably suggested it was not winter). Our participants did not have the tendency to imagine objects with similar details, or they did not remember them as often if they did. Interestingly, our participants had a greater tendency to express doubt about the accuracy of their memories more often with their real than imagined memories. Memories for imagined objects also referenced the spatial location of those objects (and others) more often than memories for real objects. This latter difference in references in spatial locations of objects may, as in the case with references to seasons, be a function of the images we selected for our real sources. That is, because the images generally portrayed close, cropped, images of the objects, there was little information on spatial locations of objects in real sources for our participants to encode. Finally, there was a greater tendency in imagined sources to reference sounds heard during the event. This difference is also likely an artefact of using pictures as stimuli for real events. Memories for imagined objects did, however, contain fewer words (M = 22.6) than descriptions of real objects (M = 25.4), $t(178) = 2.2$, $p = .03$. It should be noted that these significant differences are all rather small and generally suggest an overall lack of differences between our memories for real and imagined objects. We expected greater differences in references to sensory attributes and cognitive operations, but these differences were not obtained. It should be noted, however, that the orienting task (Please remember the photograph or your imagined picture in as much detail as possible) may have resulted in greater sensory details for generated items than might be expected under more unintentional encoding conditions. Furthermore, writing about imagined events may make them more concrete, resulting in fewer references to cognitive operations (Johnson, 1988). Either of these two factors may have contributed to fewer differences between the two classes of description.[2] Given the relative lack of differences between the classes of descriptions, we were particularly interested in whether participants would still be able to discriminate between descriptions of real and imagined objects.

Procedure

Thirty new participants were told about the nature of the stimuli preparation that had taken place prior to their participation. They were informed that they would read the descriptions of objects that had either been seen or imagined and that although any obvious references to the original source had been removed, the descriptions did not reflect any overt attempt at deception. Software randomised the presentation of all 180 descriptions (90 of each source) anew for each participant. Without any further guidance from the experimenter, they were asked to identify these descriptions as either *real* or *imagined* by pressing one of two keys on the keyboard denoting the two response options. Responses were self-paced.

[2]All descriptions were also examined using the Linguistic Inquiry and Word Count software (LIWC; Pennebaker, Booth, & Francis, 2007). The LIWC counts the number of times a written text contains self-references, social words, emotion words, cognitive words, long words (>6 letters), and articles. Analyses of our real and imagined memories revealed no significant differences across any of these dimensions. The lack of differences in the dimensions measured by LIWC is consistent with the JMCQ analysis. We thank an anonymous reviewer for suggesting this analysis.

TABLE 1
Mean memory ratings on each Judgments of Memory Characteristics Questionnaire (JMCQ) item

			Origin	
Criterion	*Scale*	*Item*	*Real*	*Imagined*
Clarity/vividness	Low–high	Clarity	1.56	1.60
		Quantity of visual details	1.39	1.42
		Vividness	1.12	1.16
		Precision of details	1.12	1.23
		Order of events	0.01	0.05
Sensory information	None–many	Colours	1.00	0.92
		Sounds	0.00	<0.04
		Smells	0.01	0.00
		Touch	0.07	0.04
		Taste	0.02	0.02
Spatial information	Not precise–precise	Location	0.43	0.53
		Setting	0.69	0.78
		Locations of objects	0.54	<0.74
		Locations of people	0.01	0.04
Indications of time	Not precise–precise	Time	0.14	0.07
		Year	0.00	0.00
		Season	0.28	>0.11
		Day	0.01	0.01
		Hour	0.01	0.05
		Age of storyteller	0.01	0.01
		Duration	0.00	0.00
Emotions and feelings	None–many	Evaluation/tone	0.04	0.03
		Role-played	0.04	0.03
		Remembered feelings	0.04	0.04
		Type of feelings	0.04	0.03
		Intensity of feelings at time	0.03	0.02
		Intensity of feelings now	0.03	0.02
		Implications to personality	0.02	0.02
Story reconstructability	Not precise–precise	Complexity	1.12	1.01
		Presumed Consequences	0.00	0.00
		Factual consequences	0.00	0.00
		Quality of remembering	1.53	1.54
		Description of events before	0.03	0.02
		Description of events after	0.00	0.00
		Doubts about accuracy	0.09	>0.01
Realism	Not realistic–realistic	Realism	1.95	1.91
		Likelihood	1.89	1.82
		Extraordinariness	0.01	0.01
		Incredible details	0.56	0.64
		Believability	1.97	1.92
Cognitive operations	None–many	Thoughts	0.17	0.14
		Repeated thinking about	0.00	0.00
		Repeated talking about	0.00	0.00

Due to the inflation of Type I error with multiple *t* tests, only differences that were significant at .01 are marked above. Several additional dimensions differed at marginal levels (i.e., location of event, location of objects, time of event, hour of event, likelihood of event and believability of event), ps $< .09$.

Results and discussion

Our primary interest was whether participants could accurately identify the source of another person's memory as real or imagined using only a written description of the memory. Every trial consisted of two response options (i.e., real or imagined), leaving chance performance at 50%. Participants accurately identified the correct source of the descriptions above chance ($M = .60$), $t(29) = 7.96$, $p < .001$, demonstrating that people can monitor the source of another person's written memory.[3]

[3] It is possible that some percentage of items possibly referenced the item's status. For example, the words *remember*, *appear*, and *picture* occurred in 22 seen but only 6 imagined descriptions. Removing those items does not qualitatively change the pattern of results presented herein.

A report by Vrij, Edward, Roberts, and Bull (2000) found that people are more accurate in identifying truthful (67%) than imagined events (44%). The same trend emerged for real (63%) and imagined (57%) events in our study, $t(29) = 1.956$, $p = .06$. Furthermore, examining accuracy for individual participants allowed us to determine that the proportion of our participants with accuracy rates over chance (.90, chance = .50) was far greater than the proportion below chance (.10), $\chi^2(1, N = 30) = 19.2$, $p < .01$. Although average performance across participants did exceed chance levels, it could be said that the ability to reality monitor another person's memory is modest, at best. When a similar procedure has been used with oral descriptions of real and imagined events, discrimination was near chance (approximately 54% accuracy with *unrehearsed* memories, Johnson & Suengas, 1989). It is surprising, therefore, that participants in our study were able to accurately discriminate between the two classes of description. Our next question was whether the ability to reality monitor another person's memory could be affected through training (Experiment 2) and feedback (Experiment 3).

Some research has focused on training people to discriminate real from imagined events in the context of eyewitness testimony, usually involving deception (Masip et al., 2005). Generally speaking, participants trained with explicit reality-monitoring criteria (discussed in Masip et al., 2005) are better than untrained participants at discriminating between truthful and fabricated events. In fact, significant efforts have been made to establish a process by which eyewitness testimony can be validated using subtle linguistic and non-linguistic (e.g., physiological) cues (e.g., Podlesny & Raskin, 1977; Undeutsch, 1989). We hope to add to that already rich literature by providing a simple method for improving IRM accuracy. It is important to note that in most of these studies people believed the descriptions to contain explicit attempts at deception. The experiments that follow differ from this approach in two important ways. Our participants were told that there was no deception in any of the descriptions. That is, the descriptions of both real and imagined objects were truthful and represented faithful descriptions of the original participants' memories. We chose to eliminate deception as a factor in our research because its presence has been shown to influence *how information* is used when deciding that a memory was actually experienced (Johnson et al., 1998). Participants in the Johnson et al.'s study who suspected deception were more likely to consider additional sensory details as evidence that the original memory was imagined and not perceived compared with a low-suspicion control group. Accordingly, we wanted our participants to evaluate the descriptions without any suspected deceit, as they would when reality monitoring their own memory output. We also chose to use a more implicit style of training in our experiments. None of our participants were familiar with the characteristics emphasised by the Memory Characteristics Questionnaire. Instead, we were interested in whether experience with examples of each class of description (accompanied by accurate or inaccurate labels denoting the class/origin of the description) could affect discrimination accuracy (Experiment 2). We also included a mistraining condition to evaluate the stability of participants' baseline performance. That is, we were interested in whether a mistraining manipulation would be powerful enough to reduce performance to near chance.

EXPERIMENT 2: (MIS)TRAINING

Methods

Participants

Ninety new participants from the University of Georgia were randomly assigned to one of three between-subjects conditions (30 in each). In addition to training and mistraining conditions, a control condition was included that was identical to Experiment 1.

Methods

The same descriptions from Experiment 1 were used in Experiments 2 and 3. In the training condition, participants were shown 60 examples (30 from each source) of the type of descriptions they would be later judging as either seen or imagined. Labels denoting the correct source appeared beneath each description. The mistraining condition was identical to the training condition except that the label denoting source was always incorrect (i.e., descriptions of real objects were labelled as imagined). Which description appeared in the (mis)training phase was randomised anew for each participant. This phase was self-paced so that participants could spend as much time as they needed assimilating both the source and the description itself. When this phase was over they continued into a monitoring procedure identical to Experiment 1, where they were asked to judge the source of the remaining 120 descriptions (60 from each source). The control condition was identical to Experiment 1 except that participants only judged 120 descriptions (60 from each source).

Results and discussion

As in the previous experiment, the control condition was able to correctly identify the source (real or imagined) of the descriptions above chance ($M = .60$, $SD = .06$). Furthermore, training improved discrimination accuracy ($M = .65$, $SD = .08$) and mistraining reduced performance to near-chance ($M = .53$, $SD = .08$). The results of an analysis of variance (ANOVA) indicated that (mis)training significantly affected discrimination accuracy, $F(2,87) = 17.81$, $p < .001$, $\eta_p^2 = .29$. Post hoc comparisons using Fisher's least significant difference test revealed that all three conditions were significantly different from each other, all $ps < .025$.

The results from Experiment 2 demonstrate that it is possible to use (mis)training to positively and negatively affect people's ability to reality monitor the source of others' memory. The (mis) training our participants received in Experiment 2 may have emphasised subtle cues present in a given description thereby sharpening participants' metamnemonic concepts of real versus imagined memories that could be applied to future discriminations. To investigate situations in which we cannot hone our judgments until after we have made them, Experiment 2B used a feedback system whereby participants were informed of their accuracy only after they had made a judgement. The baseline performance from Experiment 1 suggests that participants will make the correct judgement on 60% of trials, and thus, receive positive feedback. On the remainder of trials on which they are incorrect, they will have chosen the wrong source for the description before receiving negative feedback. Negative feedback has been shown to be less effective than positive feedback at evincing behavioural changes (Hattie & Timperley, 2007) and in some cases feedback only results in changes to a final (as opposed to an immediate) test (Smith & Kimball, 2010). To investigate these ideas, we examined whether varying the number of trials on which participants received feedback would determine the presence and extent of any improvement (Experiment 2B).

EXPERIMENT 3: FEEDBACK

Because we believed a feedback procedure would be less effective (although perhaps more ecologically valid) than training, we included two feedback conditions: FB60 provided feedback to participants on the first 60 trials and FB180 provided feedback on all 180 trials. If feedback is as effective as training, then participants should show the same benefit after 60 trials of feedback as they did after 60 trials of training. By contrast, if feedback is less effective than training then it may require feedback on far more trials before any improvement is detected.

Methods

Participants

Ninety participants from the University of Georgia participated in this experiment for partial fulfilment of a course research requirement. Participants were randomly assigned to one of three between-subjects conditions (30 in each). In addition to two conditions with varying amounts of feedback, a control condition was included that was identical to Experiment 1.

Methods

Participants judged all 180 descriptions as either *real* or *imagined*. All participants were informed that they would receive immediate feedback on their judgments. Participants in the FB60 condition were given accurate feedback on a blank screen (e.g., "You were correct, that was Imagined") immediately after making their judgement for the first 60 descriptions while participants in the FB180 condition received feedback throughout the experiment (i.e., all 180 trials). As before, all reality-monitoring judgments were self-paced and participants could choose to view their feedback (not the original description) for as long as they wished before proceeding to the next description.

Results and discussion

Including the first 60 trials of feedback into our measure of reality monitoring may obscure any benefit participants experienced in the FB60 condition because it would include early trials on which any potential benefit might not have occurred. Including these trials, however, did not affect the overall pattern of results, so they were included in subsequent analyses. An ANOVA and subsequent post hoc comparisons on reality-monitoring performance as a function of condition revealed that feedback improves performance, but only if the participants received feedback on every trial, $F(2,87) = 4.67$, $p = .01$, $\eta_p^2 = .085$. Results from Fisher's LSD test revealed that participants in the FB180 condition ($M = .64$) improved their performance above baseline and the FB60 condition, both ps < .01, while the FB60 condition ($M = .60$) was not different from the control ($M = .59$). It is possible that the participants in the FB180 condition approached the task differently from participants in the FB60 condition as a result of knowing how much feedback they would be receiving. In order to rule out any suspicion that demand characteristics were the cause of the improvement and not the continued feedback itself, we compared the two feedback conditions on the first 60 trials, during which all participants received feedback. Participants in both conditions demonstrated no improvement above baseline and no differences between each other after the first 60 trials (both Ms = .60), thus the improvement in overall performance evident in the FB180 condition was somewhat a function of the greater amount of feedback they experienced throughout the experiment. Interestingly, a recent review of feedback effects reported smaller effect sizes for less frequent feedback compared with more frequent feedback (Hattie & Timperley, 2007), however the extent to which other factors may have modified the effect sizes is unknown.

GENERAL DISCUSSION

Participants in our experiments consistently demonstrated an ability to discriminate real from imagined events relying on only the written descriptions of those events. Moreover, their ability to discriminate real from imagined memories was amenable to training, mistraining and feedback. These results are particularly intriguing because qualitative differences typically found between memories for real and imagined events were generally not present in the written descriptions of these memories (i.e., the majority of SMCQ characteristics were equivalent across both real and imagined descriptions). As discussed earlier, it is possible that writing about an imagined memory causes it to become more like a memory for something actually real (Johnson, 1988). Under the best of circumstances, the ability to discriminate was weak, however, it is impressive that discrimination obtained given the relative lack of rated differences between the descriptions. It is certainly possible that the descriptions differed on a dimension not captured during our rating process that participants exploited in order to make accurate judgments. Although the current results are curious, they only begin to explain how participants made their judgments. Future experimental work investigating the characteristics and decision processes that aid in resolving the source of others' memories will increase the applicability of these findings.

Building on this idea, the current experiments held constant the descriptions of memories for real and imagined events (i.e., the characteristics of the events were unchanging across all three experiments). Thus, our manipulation of training, mistraining and feedback influenced participants' approach towards making their decisions but not the material that they were using to make their decisions. By selectively influencing performance, we have demonstrated that an individual's ability to reliably discriminate descriptions of others' memories is dependent on evaluation and decision processes. These findings are closely related to Johnson and colleagues notion of a mental agenda that people use when they make source discriminations (see Mitchell et al., 2008). Presumably, participants in the current study changed their mental agenda as a function of training, mistraining and feedback. The mental agenda that an individual instantiates to make a source decision specifies which characteristics should be weighted most heavily in their source decisions (i.e., perceptual or cognitive operations). Additionally, the mental agenda determines other decision processes such as whether or not a liberal criterion is acceptable for making decisions. Future theoretical and experimental work is needed to determine what other variables influence participants' mental agenda used to judge others' memories (e.g., working memory, payoffs, time pressure, etc.). It would also be interesting to know the longevity of any changes to IRM that we made with our training and feedback procedures. That is, for how long after the study were our participants better (or worse) at IRM judgments?

A better understanding of IRM would have implications for theorising in different domains including eyewitness testimony, confabulation, source-monitoring and detecting conscious and unconscious plagiarisms. Employing principles of the source-monitoring framework to applied psychological issues has been useful for understanding behaviour (e.g., lie detection; Masip et al., 2005). In a similar way, the source-monitoring framework provides hypotheses about the mechanisms involved in how we analyse the characteristics of others' memories to infer how they acquired that information.

REFERENCES

Barnier, A. J., Sharman, S. J., McKay, L., & Sporer, S. L. (2005). Discriminating adults' genuine, imagined, and deceptive accounts of positive and negative childhood events. *Applied Cognitive Psychology*, *19*, 985–1001. doi:10.1002/acp.1139

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, *56*, 669–679. doi:10.1037/0022-3514.56.5.669

Bower, G. (1967). A multicomponent theory of the memory trace. In W. S. Kenneth & S. Janet Taylor (Eds.), *Psychology of learning and motivation* (vol. 1, pp. 229–325): Oxford: Academic Press.

Finke, R. A., Johnson, M. K., & Shyi, G. C.-W. (1988). Memory confusions for real and imagined completions of symmetrical visual patterns. *Memory & Cognition*, *16*(2), 133–137. doi:10.3758/BF03213481

Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, *54*(1), 51–60. doi:10.2307/1129860

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging*, *5*(1), 119–126. doi:10.1037/0882-7974.5.1.119

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

Henkel, L. A., Franklin, N., & Johnson, M. K. (2000). Cross-modal source monitoring confusions between perceived and imagined events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 321–335. doi:10.1037/0278-7393.26.2.321

Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 503–508. doi:10.1037/0278-7393.28.3.503

Hyman Jr., I. E., & Loftus, E. F. (2002). False childhood memories and eyewitness memory errors. In M. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview: Personality and clinical psychology series* (pp. 63–84). Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*, 390–394. doi:10.1037/0096-3445.117.4.390

Johnson, M. K., Bush, J. G., & Mitchell, K. J. (1998). Interpersonal reality monitoring: Judging the sources of other people's memories. *Social Cognition*, *16*, 199–224. doi:10.1521/soco.1998.16.2.199

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. doi:10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. doi:10.1037/0033-295X.88.1.67

Johnson, M. K., & Suengas, A. G. (1989). Reality monitoring judgments of other people's memories. *Bulletin of the Psychonomic Society*, *27*(2), 107–110. doi:10.3758/BF03329910

Keogh, L., & Markham, R. (1998). Judgements of other people's memory reports: Differences in reports as a function of imagery vividness. *Applied Cognitive Psychology*, *12*, 159–171. doi:10.1002/(SICI)1099-0720(199804)12:2<159::AID-ACP506>3.0.CO;2-J

Lyle, K., & Johnson, M. (2006). Importing perceived features into false memories. *Memory*, *14*, 197–213. doi:10.1080/09658210544000060

Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime and Law*, *11*(1), 99–122. doi:10.1080/10683160410001726356

Mazzoni, G., Scoboria, A., & Harvey, L., (2010). Nonbelieved memories. *Psychological Science*, *21*, 1334–1340. doi:10.1177/0956797610379865

Mill, J. S. (1869/1967). *Analysis of the phenomena of the human mind, by James Mill* (edited with additional notes by J. S. Mill). New York, NY: Kelley.

Mitchell, K. J., Raye, C. L., McGuire, J. T., Frankel, H., Greene, E. J., & Johnson, M. K. (2008). Neuroimaging evidence for agenda-dependent monitoring of different features during short-term source memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 780–790. doi:10.1037/0278-7393.34.4.780

Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginations, and 'false memories.' *Applied Cognitive Psychology*, *16*(2), 125–134. doi:10.1002/acp.779

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. Austin, TX: LIWC.net.

Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. *Psychological Bulletin*, *84*, 782–799. doi:10.1037/0033-2909.84.4.782

Schooler, J. W., Clark, C., & Loftus, E. F. (1988). Knowing when memory is real. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (vol. 1, pp. 83–88). Chichester: Wiley.

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 171–181. doi:10.1037/0278-7393.12.2.171

Smith, T. A., & Kimball, D. R. (2010). Learning from feedback: Spacing and the delay-retention effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 80–95. doi:10.1037/a0017407

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, *11*, 373–397. doi:10.1002/(SICI)1099-0720(199710)11:5<373::AID-ACP461>3.0.CO;2-0

Sporer, S. L. (2004). Reality monitoring and the detection of deception. In P.-A. Granhag & L. Stromwall (Eds.), *Deception detection in forensic contexts* (pp. 64–102). Cambridge: Cambridge University Press.

Sporer, S. L., & Kuepper, B. (2004). Fantasy and reality-memory characteristics of self-experienced and invented stories. *Zeitschrift fur Psychologie*, *212*, 135–151. doi:10.1026/0044-3409.212.3.135

Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology*, *20*, 837–854. doi:10.1002/acp.1234

Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. *Journal of Experimental Psychology: General*, *117*, 377–389. doi:10.1037/0096-3445.117.4.377

Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101–119). Dordrecht: Kluwer.

Vrij, A. (2000). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, & Law*, *11*, 3–41. doi:10.1037/1076-8971.11.1.3

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, *24*, 239–263. doi:10.1023/A:1006610329284