Differentiating Accounts of Actual, Suggested and Fabricated Childhood Events Using the Judgment of Memory Characteristics Questionnaire

JENNIFER L. SHORT and GLEN E. BODNER*
Department of Psychology, University of Calgary, Calgary, Canada

Summary: Statement analysis procedures are used in forensic settings to classify reported events as experienced or non-experienced. These procedures are typically validated using accounts of actual events and intentionally fabricated events. However, people can also unintentionally develop false memories. To examine whether inclusion of accounts of suggested events affects classification accuracy, we validated the judgment of memory characteristics questionnaire (JMCQ) statement analysis procedure using all three statement types. Participants attempted to recall two actual events and one suggested event from their childhood over two cognitive interviews, then intentionally fabricated an account of another childhood event. Fourteen of the 34 participants (41%) reported having experienced the suggested event. Independent raters then used the JMCQ to analyse and classify each type of statement from this participant subset. Inclusion of accounts of suggested events did not reduce classification accuracy. Raters tended to classify accounts of both fabricated and suggested events as non-experienced. Copyright © 2010 John Wiley & Sons, Ltd.

INTRODUCTION

People sometimes intentionally lie about events they experienced. In forensic situations, fabricated reports can impede the course of legal investigations and compromise the outcomes. Statement analysis procedures have been developed to help investigators distinguish between accounts of actual versus fabricated events by examining their qualitative characteristics (e.g. the presence vs. absence of perceptual details). One such procedure, criteria-based content analysis (CBCA), was developed based on the Undeutsch hypothesis—that there are phenomenological differences between statements based on experience versus imagination (Undeutsch, 1989; Vrij, Edward, Roberts, & Bull, 2000). CBCA has been used extensively to help researchers attempt to determine the truthfulness of forensically relevant statements.

However, in addition to reports of actual or intentionally fabricated events, people sometimes report events that were suggested to them but which they did not experience. The reality-monitoring framework (Johnson, Foley, Suengas, & Raye, 1988; Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981) proposes that internal memories (e.g. thoughts and imaginings) are sometimes mistaken for external ones. Statements based on internal events may contain more references to cognitive operations (e.g. repeatedly thinking about the event) than statements based on external events, because imaginal processes are less automatic than perceptual processes. In contrast, statements based on external events should contain more references to spatial, temporal, sensory and semantic details than statements based on internal events because perception provides more salient contextual and perceptual information than imagination. Sporer developed a statement analysis procedure based on this framework, the judgment of memory characteristics questionnaire (JMCQ; Sporer, 2004), in which ratings of the presence versus absence of different characteristics in statements is used to determine statement credibility. During recall, details from external and internal memories can also be combined (e.g. Johnson & Raye, 1981), resulting in source-monitoring errors wherein an internal memory is unintentionally believed to have been experienced, or vice versa.

Only one study to date has examined whether statement analysis can be used to distinguish accounts of actual, suggested and fabricated events. Blandon-Gitlin, Pezdek, Lindsay, and Hagen (2009, Experiment 2) had participants perform various activities in the lab (e.g. sticking pins in dolls and taking photos of bugs). A week later, they attempted to recall details about one activity of each type. Independent raters then rated these free recall statements using the CBCA. Overall CBCA scores were higher for actual events than for fabricated events, and were marginally higher for fabricated events than for suggested events. However, the CBCA was better able to differentiate the account types when the suggested event recall was deemed a 'full' memory rather than a 'partial' memory. Statement analysis may thus be useful for differentiating these three accounts, but its success will depend on the specific properties of the accounts. Building on Blandon-Gitlin et al. (2009), our study examined whether trained raters might be able to use the JMCQ to differentiate these three statement types. We chose the JMCQ rather than CBCA because it is grounded in basic psychological theory (see Sporer, 1997, 2004; Sporer & Sharan, 2006) and may be easier to code than CBCA (Sporer, 1997). We chose to use autobiographical childhood events rather than lab-based activities to increase ecological validity. The family-member-confederate paradigm was used to obtain statements for actual and suggested childhood events (e.g. Loftus & Pickrell, 1995). In this paradigm, actual childhood events are obtained from parents and are then presented to participants, usually via a title and a brief description or a set of details. A suggested childhood event that the participant did not experience is also

*Correspondence to: Glen E. Bodner, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4. E-mail: bodner@ucalgary.ca

Copyright © 2010 John Wiley & Sons, Ltd.
presented. Participants attempt to recall each event over two or three interviews. Across studies, anywhere from 0% (Pezdek, Finger, & Hodge, 1997) to 82% (Garry & Wade, 2005) of participants report the suggested event. We used plausible suggested events to help increase the rate of suggested event reports (e.g. Pezdek et al., 1997).

We also used the enhanced cognitive interview (CI; Fisher & Geiselman, 1992) rather than the free recall used by Blandon-Giltn et al. (2009) in part because statements obtained via free recall may not provide enough detail for effectively applying statement analysis. In addition, the CI was developed for questioning witnesses in forensic settings (Fisher & Geiselman). The CI begins with a free recall phase, and the interviewer then probes each reported detail using open-ended questions while avoiding leading questions.

Forensic investigations require interview procedures that will not hamper the accuracy of subsequent statement analysis. Koehnken, Schimossek, Aschermann, and Hofer (1995) found a higher rate of correct classification of statements as experienced or non-experienced using CBCA when statements were obtained with the CI rather than a structured interview. Larsson and Granhag (2005) found that statements for actual and fabricated events obtained using the CI differed on several reality-monitoring framework criteria. Statement analysis can thus be useful for distinguishing reports of actual and fabricated statements obtained with the CI.

Although the CI does not seem to impair the efficacy of statement analysis, the use of guided imagery in these interviews might promote source-monitoring errors, leading participants to misattribute details from experienced events to the suggested event. Indeed, some previous research has shown that imagery instructions can foster reports of suggested events (e.g. Hyman & Pentland, 1996). Our study is the first to measure the rate of reporting of suggested childhood events when the CI is used.

Once the cognitive interviews were conducted we trained raters to apply JMCQ statement analysis (Sporer, 2004) in which raters examine the statements for the degree to which each of 43 JMCQ items (e.g. clarity, colours and realism) are present/absent, and then provide a rating on a 7-point scale. The ratings on subsets of items are then averaged to form eight subscale ratings (see below) that raters use to try to classify statements as experienced or non-experienced.

The reality-monitoring framework predicts that statements describing actual events will receive higher ratings than statements describing fabricated events on each subscale except cognitive operations, on which fabricated statements should receive higher ratings. To date, correct classification rates of actual and fabricated statements using the JMCQ have ranged from 63 to 86% (Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004; Stromwall, Bengtsson, Leander, & Granhag, 2004; Vrij, Akehurst, Soukara, & Bull, 2004; Vrij et al., 2000), depending in part on whether statements were classified by raters or by statistical analysis.

Although correct classification using the JMCQ has not been evaluated when suggested events are included, previous research has revealed differences between accounts of actual and suggested events. Johnson et al. (1988) found that reports of actual childhood events contained higher ratings on items related to taste, realism and year of occurrence than reports of imagined childhood events. Porter, Yuille, and Lehman (1999) found that accounts of actual events were more coherent than accounts of suggested events. And Heaps and Nash (2001) found that more details, idea units and information regarding consequences were reported for actual events than suggested events.

Our study examined whether diagnostic differences between the three statement types might also emerge on the JMCQ subscales. Two raters trained to use the JMCQ attempted to classify a selected set of the childhood event statements we obtained using the CI as experienced or non-experienced. We also conducted ANOVAs to examine whether there were differences on each JMCQ subscale between statements for: (i) experienced (actual) versus non-experienced (suggested and fabricated) events; and (ii) actual versus suggested fabricated events.

The JMCQ was developed primarily using statements for actual and fabricated events, thus a central goal of our study was to determine raters’ ability to classify statements for suggested events as non-experienced. One possibility is that JMCQ ratings might be similar for suggested and fabricated event statements because both are based on internal memories. If so, raters should correctly classify both as non-experienced. Alternatively, suggested event statements might contain details misappropriated from prior actual events, and hence might be similar to actual event statements. If so, raters’ correct classification rates should be reduced when classifications for suggested event statements are included. A third possibility is that suggested event statements might contain characteristics of actual and fabricated events.

METHOD

Participants

Undergraduate psychology students (28 females, 6 males; median age of 20) from the University of Calgary participated in two interviews in exchange for course credit.

Procedure

Obtaining participants and events

Prospective participants attended an information session and were told that the study would examine adults’ abilities to recall childhood events. They were informed that the researcher needed to contact a parent to obtain descriptions of events from their childhood. For those who opted to participate, a parent was mailed a package of research materials. Of the 47 parents who received a package, 38 (81%) completed and returned it, providing titles (e.g. babysat a baby duck from school) and specific details (e.g. duck was named Sally; picked up duck from Jim; babysat duck for one day; kept duck confined to the kitchen) for 5 events experienced by their child between age 5 and 8. The materials provided were usable for 34 of these 38 participants (89%), sometimes after contacting a parent for additional information. Parents provided details for one
event they expected to be easy for their child to recall, and for four events they expected their child would have difficulty recalling. These events formed the pool of potential actual events for their child. Each parent was then phoned to help select and personalize four potential suggested and fabricated events for their child. The events provided by other parents were presented (without identifying information) and the parents selected and then personalized with plausible details four potential suggested events that they believed their child had not experienced.

Cognitive interviews
At the beginning of the first of two interviews (conducted by the first author), participants were given a form consisting of a descriptive title for each of the nine events. The four non-experienced events were presented in positions 1, 4, 6 and 9. Participants rated each event on ‘expected ease of recall’ using a 6-point Likert scale (1 = very difficult and 6 = very easy). The interviewer used these ratings to select a suggested event and a fabricated event that were rated as high as possible while also being similar in rating to the selected hard actual event. The actual event receiving the highest rating served as the easy actual event.

In the first interview, participants were presented with the event title, their age at the time, and the four details for the easy actual event, hard actual event and suggested event, then were interviewed about their memory for each event in this order. The easy actual event was presented first to allow them to become familiar with the interview procedure. The hard actual event was presented before the suggested event to lead participants to develop an expectation that some events would be hard to recall and to minimize suspicion about the suggested event.

The CI began with the interviewer building rapport and communicating the expectation that the participant would take the central role in the interview. Participants were instructed to describe everything that came to mind, without editing for potentially contradictory details or trivial information. Following rapport building, the information-gathering phase began. Participants were asked to form an image of the physical environment related to the event and to recall their state of mind at the time. They then provided a free recall of the event without interruption from the interviewer. The interviewer probed each mental image for more information by restating the participant’s description of the image. Participants were then directed to particular aspects of each mental image and were asked open-ended questions about them. This probing continued until all aspects of each mental image had been examined. To complete the interview, the researcher repeated the provided information and gave participants an opportunity for corrections and/or additions. Participants were instructed to continue thinking about each event until the second interview but to not discuss them with friends or family.

A second interview following the same procedure took place a week later, and provided an opportunity for participants to supply more detailed recall of each event. The entire second interview formed the statement that was later transcribed and analysed. Following this second round of attempted recall, participants were informed that they had not experienced one of the events. They were then asked to pick the non-experienced event and to rate their confidence in their event choice, as well as their confidence that they had not experienced this event (1 = not at all confident and 7 = completely confident).

Using the same procedure, participants were then asked to fabricate a statement for a fourth event that they had not experienced. They were asked to provide a statement that was as realistic as possible, but to not base the statement on an event from their life. The fabricated event was presented only at the end of the second interview to reduce suspicion about the suggested event.

Classifying recall of the suggested event
To determine the rate of recall of the actual, suggested and fabricated event statements, two independent judges were trained using Hyman and Pentland’s (1996) criteria to classify the recall of each event as either: (i) clear recall (event was recalled and clear elaborations were provided); (ii) partial recall (event was not recalled but consistent elaborations were provided); (iii) attempted recall (event was not recalled and elaborations were not provided, but related information or self-knowledge was conveyed); or (iv) no recall. Only statements scored as clear or partial recall by both judges were classified as recalled. Lindsay, Hagen, Read, Wade, and Garry (2004) recommended a classification system that distinguishes ‘memories’ from ‘images but no memories’. The former is equivalent to clear recall, but the latter would include a mixture of partial recall and attempted recall. We did not wish to consider attempted recalls, thus we used the Hyman and Pentland criteria.

JMCQ training and analyses
The JMCQ consists of 43 items on which statements are rated using 7-point Likert scales. Unique subsets of these items are then combined to form eight subscales (see Table 2; Sporer, 2004). Correcting an error in Sporer (2004), the cognitive operations subscale used items 31, 34, 35, 37 and 38.

The first author trained two independent raters to identify the characteristic components of each transcribed statement using the JMCQ criteria. The raters were blind to the study’s purposes, and were only informed that they would be rating statements for both experienced and non-experienced events. The raters first practiced rating Blandon-Gitlin et al.’s (2009, Experiment 2) actual and fabricated statements. Once raters were providing similar item ratings for these statements (generally within 1 point range), the actual and fabricated event statements provided by the participants who did not recall the suggested event in our study were rated to provide experience analysing interview statements. Ratings for the items that contribute to each JMCQ subscale were then averaged to obtain eight subscale ratings for each of these statements.

Based on the range of scores obtained for these statements, raters subjectively decided whether a participant’s score on each subscale was low or high. They then compared the characteristics of each statement to the expected patterns for experienced and non-experienced statements. Specifically, non-experienced statements were expected to have high
ratings on cognitive operations while experienced statements were expected to receive high ratings on the other subscales (Johnson & Raye, 1981; Johnson et al., 1993; Sporer, 2004).

Based on the subscale scores and these criteria, the raters then classified the event as experienced or non-experienced. When the raters disagreed on the classification of a statement the guidelines were discussed by both raters and the first author. Raters continued to rate subsets of these practice statements until they accurately classified all of them.

Once trained, the raters used the JMCQ to rate the actual, suggested and fabricated event statements obtained from the second interview from the subset of participants who recalled the suggested event. Based on their individual subscale ratings, each rater then attempted to classify each statement as experienced or non-experienced.

**RESULTS AND DISCUSSION**

We report our findings in two sections. The first reports the rate of recall of the suggested event using the CI, and the second reports JMCQ rater classification accuracies and subsequent ANOVAs of the JMCQ subscale scores. A significance level of .05 was used for statistical tests except where otherwise noted. Partial eta-squared ($\eta_p^2$) is reported as a measure of effect size.

**Recall of the suggested event**

The two judges classified each statement as clear recall, partial recall, attempted recall or no recall. Inter-rater reliability for classification of the actual, suggested and fabricated event statements for the 34 participants using the Hyman and Pentland (1996) categories was calculated by correlating these categorical classifications; the resulting Spearman correlation was .91. The few disagreements were resolved through discussion. Our use of the CI resulted in a rate of suggested event reporting that was higher than the 33% average reported across nine narrative-based studies reviewed by Garry and Gerrie (2005). In our study, 14 of the 34 participants (41%) recalled the suggested event in the second interview (7 clear recall and 7 partial recall). Suggested events that were recalled included a dance performance, an April Fools’ Day prank, and painting a pair of jeans (see Table 1 for examples). A further 14 statements (41%) were classified as attempted recall and were not analysed further. The second-interview statements for the actual and fabricated events of these 14 participants were also classified. All 14 recalled the easy actual event (13 clear and 1 partial) and all but one also recalled the hard actual event (11 clear and 2 partial; hence there was a missing case in some repeated-measures analyses). All 14 fabricated event statements were classified as clear recall.

When asked to pick the non-experienced event, participants who recalled the suggested event were more likely than those who did not to pick an actual event ($50 vs. 0%$), $\chi^2 (df = 3, N = 34) = 12.59$. In addition, participants who recalled (vs. did not recall) the suggested event were less confident that they had correctly identified it ($M = 4.79, SD = 1.63$ vs. $M = 6.05, SD = 0.83$), $t(32) = 2.68, SE = 0.47$, and that they had not previously experienced it ($M = 4.29, SD = 1.27$ vs. $M = 6.10, SD = 1.07$), $t(32) = 4.51, SE = 0.40$. Thus, some of our participants appear to have formed compelling memories of the suggested events.

One reason for our high rate of suggested event recall was that each suggested event was another participant’s actual event, and each was personalized for the participant with the help of that participant’s parent. In addition, using participants’ expected-ease-of-recall ratings to select a

| Table 1. Excerpts from suggested event statements judged as clear, partial or attempted recall |
|----------------------------------|----------------------------------|----------------------------------|
| **Clear recall**                | **Partial recall**               | **Attempted recall**             |
| **Event title:** Sleepover at B.’s | **Provided event details:** Painted t-shirts at the sleepover; stayed up until 1 a.m.; had pancakes for breakfast; went home for church and Thanksgiving dinner | **Event title:** April Fools’ Day prank |
| **Statement excerpt:** I do remember painting the t-shirt. I remember doing that after dinner. I’m not even sure if I did get there before dinner or not, but I remember they were black. I remember that. And we used those dipping cone things; I don’t know what they’re called. But they’re fun. They still sell them; they’re hot items. So I remember painting a shirt but I think I, I did something really weird like, did like, kind of a big kind of collar around the front, like a big section of it coloured in. And then we drew things all over it. I don’t know, it was, it was a crappy shirt. | **Provided event details:** Played on Mom; changed the sugar to salt; waited outside the kitchen; Mom put salt in her coffee and was really mad | **Event title:** Easter egg hunt |
| **Statement excerpt:** I think I remember standing on like, right beside those stairs like leaning around the corner, trying to watch her without her seeing me watching. But I don’t remember changing the salt or the sugar and I don’t really remember her putting it in or anything. But I kinda remember me and my sister just standing around the corner where, to get into the kitchen and standing beside these stairs watching into the kitchen. But I don’t remember what happened or her getting mad or anything, but it seems like something I would do. Cause my sister and I were always kinda partners in crime and we were always trying to do stuff like this and so it would seem like something I would do but very vague. | **Provided event details:** At M.’s farm; walked through barns to see horses; lots of dogs were there; ate hot dogs around a campfire | **Statement excerpt:** I still can’t remember going on an Easter egg hunt. I don’t know why. I remember doing other things but I don’t know if I can’t remember, like whether I’m thinking about the wrong place or something because M., or the M. I’m thinking of, didn’t have a farm at that point. She still lived in N. and everything. So I don’t know if I’m thinking of somebody else. The farm that she did have later and everything, we did stuff out there. We’ve had scavenger hunts and everything but I don’t remember actually going on an Easter egg hunt out there. |

---

plausible suggested event likely increased the general plausibility (its likelihood for that cohort) and personal plausibility (its likelihood for that participant) of the suggested event (Scoboria, Mazzoni, Kirsch, & Relyea, 2004). Using plausible suggested events likely fostered participants’ belief that the event occurred, which may be a precursor for forming a false memory (Hyman & Loftus, 1998).

Our use of the CI rather than a simple free-recall task may also have fostered reporting of the suggested event. The CI uses both context reinstatement and guided imagery, two memory retrieval techniques that may promote unintentional false recall (e.g. Garry & Wade, 2005; Hyman & Pentland, 1996; Lindsay et al., 2004; Porter et al., 1999). Probing participants’ mental images in detail may also have activated schemas and actual memories that were integrated into their reconstruction attempt via source-monitoring errors (Johnson et al., 1993; Lyle & Johnson, 2006). Although we did not include a comparison interview procedure here, the high rate of recall of the suggested events suggests that the CI should be used with care in forensic settings.

### JMCQ analyses

We restricted our subsequent analyses to the statements provided by the 14 participants who clearly or partially reported the suggested event, for several reasons. First, this ensured that raters received a similar number of experienced (N = 27; one hard actual event was not recalled) and non-experienced (N = 28) statements. Analysis of statements from all 34 participants would have biased raters toward rating statements as experienced due to the higher proportion of such statements in the pool. Second, because the same participants provided both the experienced and non-experienced statements, any differences between statement types cannot be attributed to (i.e. are not confounded by) differences between participants who report versus do not report suggested events. Third, this procedure allowed the raters to be trained on the remaining participants’ actual and fabricated statements.

Table 2 shows raters’ average JMCQ subscale scores for statements based on experienced events (easy actual and hard actual events) and for statements based on non-experienced events (suggested and fabricated). The inter-rater reliability among the two raters’ JMCQ subscale scores (excluding the realism subscale here and in all subsequent analyses because all events were rated 7 by both raters) were as follows: clarity and vividness ($r = .67$), sensory information ($r = .70$), spatial information ($r = .68$), time information ($r = .57$), emotions and feelings ($r = .65$), reconstructability of the story ($r = .68$) and cognitive operations ($r = .41$). These reliabilities were each significant, but were also lower than anticipated, in part due to a restricted range of scores on each subscale. However, correct classification of statements using the composite statement analysis ratings can occur even when the inter-rater reliability of individual items or subscales is modest (e.g. Masip et al., 2005; Vrij, 2005).

The raters classified each statement as experienced or non-experienced based on the trends in the JMCQ subscale scores. Rater agreement on the classification of statements as experienced versus non-experienced was very good (Cohen’s kappa = 0.86). Averaged across the two raters, 74.1% of actual event statements, 67.8% of suggested event statements, and 67.8% of fabricated event statements were correctly classified. The overall correct classification rate of 70.9% is comparable to previous studies that included only actual and fabricated statements (see Sporer, 2004). Removal of the suggested event statements resulted in a correct classification rate of 72.0%. Thus, classification of events as experienced versus non-experienced was little affected by the inclusion of suggested event statements.

Next turn to analyses comparing the JMCQ subscale ratings of various statement types. To attain a single set of subscale scores for each statement type for the ANOVAs, the two raters’ scores for each of the 43 items were first averaged. The subscale scores were then determined by averaging the ratings for the items contributing to each subscale.

We first compared the ratings obtained for experienced and non-experienced event statements using separate one-way repeated-measures ANOVAs for each JMCQ subscale (excluding realism). Clarity-and-vividness subscale ratings were higher for experienced than non-experienced event statements ($M = 4.15, SD = 0.60$ vs. $M = 3.57, SD = 0.51$), $F(1, 12) = 7.24, MSE = 0.29, \eta_p^2 = 0.38$, and this difference was marginally significant for the time-information subscale.

### Table 2. Mean JMCQ subscale scores (and standard deviations) for statements based either on experienced events (easy actual and hard actual) or non-experienced events (suggested and fabricated)

<table>
<thead>
<tr>
<th>JMCQ subscale</th>
<th>Experienced</th>
<th>Non-experienced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Easy (N = 14)</td>
<td>Hard (N = 13)</td>
</tr>
<tr>
<td>Clarity and vividness</td>
<td>4.35 (.55)</td>
<td>3.98 (.91)</td>
</tr>
<tr>
<td>Sensory information</td>
<td>1.46 (.23)</td>
<td>1.41 (.30)</td>
</tr>
<tr>
<td>Spatial information</td>
<td>2.92 (.60)</td>
<td>3.00 (.79)</td>
</tr>
<tr>
<td>Time information</td>
<td>1.74 (.62)</td>
<td>1.68 (.91)</td>
</tr>
<tr>
<td>Emotions and feelings</td>
<td>2.46 (.49)</td>
<td>2.68 (.51)</td>
</tr>
<tr>
<td>Reconstructability</td>
<td>2.79 (.27)</td>
<td>2.33 (.46)</td>
</tr>
<tr>
<td>Cognitive operations</td>
<td>1.54 (.44)</td>
<td>1.35 (.33)</td>
</tr>
</tbody>
</table>

*Note*: Scores are out of 7. Higher scores indicate greater presence of that subscale characteristic.
(M = 1.74, SD = 0.52 vs. M = 1.44, SD = 0.40), F (1, 12) = 4.26, \(\text{MSE} = 0.15, p = .06, \eta^2_p = 0.26\). The differences on the remaining subscales were not significant, though with the exception of the cognitive operations subscale they were consistently in the direction predicted by the reality-monitoring framework. Raters may have found the subscale scores useful for classifying the statements even though most of the mean differences were not significant. The role of the subscale scores in raters’ classification decisions could be isolated in future studies by comparing accuracy for raters given versus not given the subscale scores.

We also conducted some exploratory analyses comparing, for each JMCQ subscale, the scores for all four types of event statement (Table 2). These one-way repeated-measures ANOVAs were significant only for the clarity-and-vividness subscale, \(F (3, 36) = 4.91, \text{MSE} = 0.48, \eta^2_p = 0.29\), and the reconstructability-of-the-story subscale, \(F (3, 36) = 9.38, \text{MSE} = 0.11, \eta^2_p = 0.44\). Follow-up tests were conducted using a Bonferroni-corrected alpha of .01. Easy actual event statements had a higher clarity-and-vividness subscore than suggested event statements (\(M = 4.35, \text{SD} = 0.55\) vs. \(M = 3.30, \text{SD} = 0.84\), \(t (13) = 4.28, SE = 0.25\); the remaining statement type differences on this subscale were not significant. On the reconstructability-of-the-story subscale, the mean for hard actual event statements (\(M = 2.33, \text{SD} = 0.46\)) and suggested event statements (\(M = 2.26, \text{SD} = 0.50\)) were each significantly lower than for easy actual event statements (\(M = 2.79, \text{SD} = 0.28\)) and fabricated event statements (\(M = 2.79, \text{SD} = 0.29\)), \(t (3) = 3.41–5.86, SEs = 0.10–0.14\). Recall of more difficult (or non-experienced) events was thus characterized by decreased reconstructability of the event. Whether these particular subscale differences between statement types are replicable remains to be seen, but their existence would mean that trained raters might have some success at using the JMCQ to distinguish reports of suggested events from reports of both actual and fabricated events.

Our participants were interviewed about the actual events and the suggested event twice (with a week to reflect in between), but about the fabricated event only once (after the last interview, as in Porter et al., 1999). Although this was done to minimize suspicion about the suggested event, differences between the fabricated statements and the other statement types may have been exaggerated as a result. In future studies, participants could be asked to recall only one type of event. Use of a between-subjects design would also allow researchers to use logistic regression to determine the optimal rate of correct classification of actual, suggested and fabricated event statements. Analyses of JMCQ subscale ratings using both binary (experienced vs. non-experienced events) and tertiary (actual vs. suggested vs. fabricated events) logistic regressions would provide another means of evaluating the utility of the JMCQ in classifying suggested event statements.

Finally, given the existence of suggested event statements, classification accuracy of the JMCQ might be improved by re-factoring the JMCQ items into subscales based on analysis of all three statement types. This re-factoring will require a larger data set than is currently available. Refactoring based on all three statement types could increase the JMCQ’s classification accuracy as well as its validity, making it more attractive for use in forensic investigations.

REFERENCES


