Assessment Criteria Indicative of Deception (ACID): An Integrated System of Investigative Interviewing and Detecting Deception

KEVIN COLWELL[1,*], CHERYL K. HISCOCK-ANISMAN[2], AMINA MEMON[3], LAURA TAYLOR[1] and JESSICA PREWETT[4]

[1]*Southern Connecticut State University, CT, USA*
[2]*National University, CA, USA*
[3]*University of Aberdeen, Aberdeen, UK*
[4]*Valdosta State University, GA, USA*

Abstract

This study describes the assimilation and validation of Assessment Criteria Indicative of Deception (ACID). ACID is derived from investigative interviewing, Criteria-Based Content Analysis, Reality Monitoring, and interpersonal deception. Each component has been previously published. Thirty-eight university undergraduates entered a professor's office and either stole an exam or replaced an exam that had been stolen previously. They were interviewed 1 week later with the Reality Interview, which is deliberately challenging and aims to enhance the detection of deception. Half responded honestly and completely; half distorted their responses to avoid incrimination. Incentives were provided. Honest responses were longer, more detailed, and contained more admissions of potential mistakes. Most importantly, honest respondents benefited from attempts to enhance recall, whereas these same attempts caused deceptive respondents to provide shorter, more repetitive statements. This is a promising technique; 33 of 38 cases were classified accurately. Discussion includes characteristics of deception, process of deception during an investigative interview, hypothetical interview strategies to facilitate the detection of deception, strengths and weaknesses of the study, and areas for future research. Copyright © 2008 John Wiley & Sons, Ltd.

Key words: investigative interviewing; deception; credibility; reality monitoring; CBCA; vividness; spontaneity

There is a crucial need to develop a valid and applicable method for obtaining and evaluating information during an investigation. The present research is a validation of one such system of interviewing and credibility assessment, known as Assessment Criteria Indicative of Deception (ACID). ACID combines content criteria derived from research in interpersonal deception and memory with investigative interviewing to facilitate the detection of deception.

*Correspondence to: Kevin Colwell, Ph.D., Department of Psychology, Southern Connecticut State University, 501 Crescent Street, New Haven, CT 06515, USA.
E-mail: colwellk2@southernct.edu

Interpersonal deception

A meta-analysis of over 1,300 estimates from 120 papers found that fabricated responses were shorter and had less supporting detail than true responses (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). Vrij and Mann (2006) found that lying is more cognitively demanding than honest reporting and that liars work harder to control their speech. In another study, these researchers found the deceptive portions of a murderer's statements to be shorter, more carefully phrased, and less spontaneous (Mann, Vrij, & Bull, 2002). Overall, these findings are consistent with Information Manipulation Theory (IMT) (McCornack, 1992), which describes deception as a balance between disclosing enough information to satisfy the interrogator or interviewer whilst protecting and hiding any information that could betray deception. Thus, the control and careful release of information becomes important to avoid detection. This control requires 'thinking hard' and constant impression management (Vrij & Mann, 2006, p. 337).

Memory and credibility assessment

The longest standing and most researched memory-based system of credibility assessment is Criteria-Based Content Analysis (CBCA) (Vrij & Mann, 2006; see Colwell, Hiscock-Anisman, Memon, Rachel, & Colwell, 2007, for a review). Two CBCA criteria, Unstructured Production (Spontaneous Reproduction) and Quantity of Details (or Sufficient Detail) are promising in distinguishing honest from deceptive statements. Taken together, these criteria posit that statements derived from genuine memory differ in amount and distribution of their detail content. Genuine memories for external events are more detailed and follow a less rigid, more spontaneous structure than memories derived from imagination or fabrication.

Another line of research that attempts to determine whether a memory is derived from genuine experience versus imagination or fabrication is Reality Monitoring (RM) (Johnson, 1988; Johnson & Raye, 1981). Like CBCA, RM decisions often are made from the amount and type of details in a statement. According to this theory, memories for genuine events should contain more external-sensorial details (e.g. colour, smell, taste, etc.) and more contextual details (e.g. temporal and spatial relationships). Conversely, imagination, contamination, and fabrication ought to involve less richness and fewer external and contextual details but more details derived from internally generated memories. Internal details are those that deal with something from the personal history of the witness or with something unique and idiosyncratic to the experience of the witness (e.g. events from the past, thoughts, opinions, etc.) (Johnson, 1988; Johnson, Foley, Suengas, & Raye, 1988; Johnson & Raye, 1981).

Interviewing and deception

Control of information and impression management make deception more cognitively demanding than honest recall. Because of this, interviewing strategies that further increase cognitive load and highlight attempts at impression management can maximise detection of deception (Vrij, Fisher, Mann, & Leal, 2006). A common strategy of deception is to prepare and rehearse a fictitious account of the target event. During questioning, the deceiver bases her or his responses on this lie script rather than recall and reconstruction of the original event. This strategy is called superficial encoding (Porter & Yuille, 1996). This relatively short, carefully phrased, well rehearsed and unchanging 'lie script' (Colwell,

Hiscock-Anisman, & Memon, 2002; Colwell, Hiscock-Anisman, Memon, Woods, & Michlik, 2006; Porter & Yuille, 1996; Vrij, 2000) allows deceivers to reduce cognitive demand and anxiety, and to appear co-operative by providing sufficient information whilst concealing any information that could lead to detection (McCornack, 1992; Porter & Yuille, 1996). In short, honest reporters try to reconstruct their memory for the target event, whereas deceptive reporters try to repeat accurately their lie script (Granhag & Strömwall, 1999; Granhag, Strömwall, & Jonsson, 2003). This difference in processing and storage leads to systematic and detectable differences in honest versus deceptive statements as they unfold across the course of an investigative interview.

The ACID system

The ACID system analyses the length of responses, admitting potential errors, and RM criteria operationalized to detect differences due to memory and impression management or control of information. This extended approach to RM tallies the unique external, contextual, and internal details as they appear during the course of an investigative interview designed to detect deception. In the first empirical test of this approach to RM, 95% of the statements of offenders who witnessed a staged theft were classified accurately as honest or deceptive. Honest statements were more vivid and spontaneous. That is: (i) honest statements had more external and contextual detail overall; and (ii) honest responders added more detail due to the recall enhancement effect of the mnemonics. The most striking and statistically powerful aspect of the study was the details that were not provided during free recall but rather were added in the course of later recall tasks (Colwell, Hiscock-Anisman, Memon, *et al.*, 2007).

The ACID approach to RM has received considerable research support (Colwell, Hiscock-Anisman, Leach, Corbett, & Uerz, 2007; Colwell, Hiscock-Anisman, Memon *et al.*, 2007; Memon, Mastroberardino, Fraser, & Colwell, 2008). These studies have included adult witnesses in college and prison, children's statements given to a Canadian Magistrate, and credibility decisions from trained raters reading transcripts. In sum, the ACID approach to RM highlights reliable, valid, and applicable differences between honest and deceptive responding.

At this point, a brief discussion of specific RM criteria is necessary. To date, there have been numerous and somewhat inconsistent rating systems derived from RM research (Masip, Sporer, Garrido, & Herrero, 2005). The current system is aligned more closely with the original postulates of RM, and there are some discrepancies between these definitions and the ones used by other authors. A notable example, Aldert Vrij's rating scheme uses visual, auditory, temporal, spatial, and cognitive operations as detail categories. Despite cosmetic differences, these two systems were derived from the same theory and correspond quite well: The original external category here contains visual and auditory information in Vrij's system, the contextual category contains temporal and spatial information, and the internal category contains cognitive operations. In a recent comparison of these two systems, the scores derived from the current system correlated between 0.70 and 0.90 with the scores derived by Vrij's system, and the two systems performed equally well in their ability to detect deception (Memon *et al.*, 2008). As such, the disagreements regarding specific detail categories appropriate for RM should not obscure the validity of the underlying concept.

The investigative interview of choice for the ACID system is the Reality Interview (RI) (Colwell *et al.*, 2002). The RI is a variation of the original Cognitive Interview (CI) (Fisher

& Geiselman, 1992; Geiselman, Fisher, MacKinnon, & Holland, 1985) that increases the cognitive demand placed on the interviewee and highlights attempts at impression management. Multiple recall tasks (some of which are designed specifically to circumvent schema-specific recall) increase demands on a deceiver who is trying to keep responses consistent with her or his rehearsed script. Two-alternative, forced-choice inferences are inserted between the mnemonic strategies of the CI. These inferences are designed to require deeper processing than simple recall of a script and, therefore, force the deceiver to think about and to reconstruct the original event. This interrupts the interviewee's use of a lie script and makes superficial encoding more difficult. In contrast, the multiple recall attempts and deeper-level processing both provide recall cues for honest reporters. Therefore, the RI exacerbates differences between honest and deceptive statements, and generates specific trends in how information is distributed throughout a statement. Honest respondents benefit from these cues, producing longer and more detailed statements as the interview progresses, whereas deceptive respondents are hindered and tend to provide shorter and less detailed statements.

The RI was compared to the Stepwise Interview (SI) and CI in a series of studies (Colwell *et al.*, 2002; Hiscock-Anisman, Colwell, Memon, & Laurita, 2008). In the initial study, credibility decisions were made using response length, type-token ratio, and coherence. Using these criteria, a higher proportion of witness statements were classified correctly as honest or deceptive using the RI than using the SI. The relative performances of the RI and CI were almost the same, with the RI yielding only a slight improvement in predictive accuracy. In the later study, which examined statement detail characteristics and response length, the RI outperformed the CI in credibility assessment. Honest reporters in the RI condition provided longer and more detailed statements. Deceptive reporters in the RI condition provided shorter and less detailed statements.[1] Thus, the RI was able to enhance the detection of deception without decreasing the amount of information obtained from co-operative witnesses (Colwell *et al.*, 2002).

Current research

The current approach has been dubbed ACID, and it derives from research in investigative interviewing, impression-management, and memory. ACID employs the RI (Colwell *et al.*, 2002) to facilitate the detection of deception by increasing cognitive load and by forcing multiple recall attempts in order to highlight efforts at impression management. Honest statements are hypothesised to be longer and more detailed, with relatively more details after the initial description of the target event. Deceptive statements are hypothesised to be shorter and less detailed, with relatively fewer details added after the initial description of the target event. Honest reporters are also more likely to draw attention to themselves by admitting potential mistakes.

In this study, the credibility assessment portion of the ACID system was compared to the Judgment of Memory Characteristics Questionnaire (JMCQ) to assess the relative utility of each approach. Another tool derived from RM research, the JMCQ has raters provide Likert-type ratings across a number of factors derived from later RM research. In previous research, the JMCQ differentiated effectively honest from imagined autobiographical memory in undergraduates (Sporer, 1997). However, the JMCQ was ineffective at discriminating honest from deceptive witness statements in the Colwell, Hiscock-Anisman, Memon, *et al.*, (2007) study. This may have resulted from the fact that all

---

[1]The accuracy of details was not assessed in this study.

participants in the latter study witnessed the same staged theft. That is, the participants in both groups reported memory for an experienced event. As such, a corollary purpose of the present study was to investigate the utility of the JMCQ with respect to credibility assessments that involve control of information rather than imagination or fabrication.

It was hypothesised that honest reports: (i) would likely contain more external and contextual and fewer internal details during the free recall phase of interview (the combination of motivated deception and passage of time may reduce or erase differences in external or contextual details); (ii) would contain more *new* external, contextual, and internal details during the mnemonic phase of the interview; (iii) would have longer response lengths for the free recall and mnemonic phases of the interview; (iv) would be more likely to contain admissions of possible errors; and (v) would receive higher scores on the scales of the JMCQ. It also was expected that the credibility assessment portion of the ACID system would outperform the JMCQ in the amount of variance explained and that the ACID technique as a whole would allow for better-than-chance classification of reports as honest or deceptive.

METHOD

Participants

Forty-three undergraduate students were recruited for this study. Five of these were dropped due to incomprehensible portions of their later recorded interview, leaving a sample of 38. Twenty-four (63%) were female and 14 (37%) were male. Sixty-eight per cent described themselves as Caucasian, 21% as African American, and 11% as Asian American.

Procedure

After meeting with a research assistant to receive the first portion of their instructions, participants entered a psychology office and either stole an exam key or replaced a stolen exam key. They were told the professor using this office did not know of this experiment and would become upset if they were discovered, requiring the intervention of the Department Chair to keep them from getting into trouble with the Campus Police. As they entered this office, participants passed an active video camera. After completing their task in the room, participants reported back to a research assistant and were provided with the second portion of their instructions, the honesty of reporting and motivation instructions.

All participants were told that they would have approximately 1 week before they would be interviewed regarding their time in the room. They were told, 'You have been seen on video entering the room, and when you return for interviewing, you must provide a convincing story that explains what you did while in the professor's office.' Honest participants were instructed to report their activities as completely as possible, whereas deceptive participants were instructed to distort their statements so they were not implicated in any wrongdoing. In an attempt to manipulate motivation, one-half of the students were told that they would receive one dollar for a successful lie and that the two best 'story tellers' would receive twenty dollars. The incentive of a prize was used to generate sufficient motivation to create a believable lie script and, thus, to mirror some of the intrinsic motivation a deceiver feels. It also was used to make deception detection more difficult.

After the week's delay, the students were interviewed using the RI regarding their activities in the room. All interviews followed the same semi-structured format (see Table 1).

Table 1.   Script for reality interview

	Recall task	Phrase from recall task	Interview portion for scoring
1.	Baseline and rapport	a. 'Last meal' b. 'First day of semester'	Not scored
2.	Free Recall	'Please describe, in as much detail as possible, everything that happened in Room 212.'	Free recall
3.	Mental reinstatement of context	'Think about and include all sights, sounds, smells, emotions, thoughts, or anything else from time of event.'	Mnemonics
4.	Inferential block 1	a. 'If a police officer had been present, would he have noticed something wrong?' b. 'Was a crime committed?' c. 'Did anyone speak with an accent?'	Not scored
5.	Recall from other perspective	'If someone else had been in the room, what would they have seen?'	Mnemonics
6.	Inferential block 2	a. 'Did anyone intend to harm anyone else?' b. 'Was this an act of violence?' c. 'Were there any weapons in the event?'	Not scored
7.	Reverse order recall	'Beginning with last, and ending with first, please describe entire event in reverse order.'	Mnemonics
8.	Inferential block 3	a. 'Did you notice anything unusual about the room?' b. 'Would anyone think that you did something you weren't supposed to whilst in the room?' c. 'Do you think that you could have been mistaken about anything you have said so far?'	'Any mistakes' scored as Yes or No
9.	Retell entire event	'Please describe, in as much detail as possible, everything that happened in Room 212.'	Mnemonics

Participants were assigned randomly to one of four interviewers, who were blind to the hypotheses, conditions, and target event of the experiment. Following the interview, participants completed a post-event questionnaire as a manipulation check of the experimental condition and overall level of motivation. The post-event questionnaire contained four Likert-type questions to assess: (i) the amount of anxiety participants experienced whilst in the room with the exam; (ii) the amount of anxiety participants experienced during the interview; (iii) the amount of anxiety participants experienced during an exam in their psychology class (for comparison purposes); and (iv) the amount of motivation they felt to tell a complete and convincing story. Each of these questions ranged from 1 (none felt) to 4 (moderate amount felt) to 7 (most extreme feeling possible). At the conclusion of the study, all interviews were transcribed and scored according to ACID and JMCQ criteria.

Materials: the Reality Interview

The RI used in this study followed the same format as in the Colwell *et al.* study (Mann *et al.*, 2002): baseline, free recall, and the multiple recall trials created by three of the traditional mnemonics (mental reinstatement of context, recall from another perspective, and reverse order recall) from the CI. The baselines used were free recall descriptions of the student's first day on campus that semester and of the last meal the student had eaten. The 'target' portion of the interview began with a free recall task. Free recall required participants to 'describe, in as much detail as possible, everything [they remembered] from the time [they] spent in the professor's office.' For mental reinstatement of context, participants were asked to 'think about and include all associated details such as sights, sounds, smells, [their] emotions, [their] thoughts at the time, or anything else [they could] remember from the time of the event.' For recall from another perspective, participants were asked to answer, 'as if someone else had been in the room.' For reverse order recall, participants were asked, 'Starting with the last detail, and ending with the first, please describe the entire event, in reverse order.' Finally, participants were asked to 'describe the entire event one more time, to make certain that I understand everything.'

Three blocks of three inferential questions each were inserted into this structure. Examples of inferential questions are, 'If a policeman had been watching, would he think a crime had been committed?,' 'Was this an act of violence?,' and 'Could you have made any mistakes in the information that you have reported so far?'[2] Table 1 displays the RI used in the current study.

Measures

Transcript ratings

Transcripts were coded by four trained raters who were blind to the conditions and the hypotheses of the experiment. All transcripts were coded for the following nine criteria: the response length and the amount of external, contextual, and internal details during free recall; the response length and the amount of external, contextual, and internal details during the mnemonics phase; and any admission of error by the respondent throughout the interview.

External details were defined as information regarding the event in question that was gained from the senses (e.g. describing who, what, and where). For example, the phrase 'a *tall man* with *red hair*' contains four external details. Contextual details describe relationships amongst objects and/or actors (e.g. temporal, spatial relationships). The sentence 'The rings were *on top of* the desk' contains one contextual detail. Internal details were defined as information regarding the subjective mood, experiences, or cognitive processes of the respondent, as well as any information that referenced the respondent's history rather than the event in question. Therefore, the sentence 'I was *nervous*' contains one internal detail.

The training of raters followed the same structure as reported in Colwell, Hiscock-Anisman, Memon, *et al.* (2007). It consisted of three one-hour group meetings, with homework assignments between two of the meetings, and a final meeting to reach consensus. During the first meeting, each of the four classes of detail was defined, and raters were provided with standard scoring sheets to ensure consistent operational definitions of the variables.

---

[2]This final question is the only one of the inferential questions that was scored. The remainder of the questions was used only to increase cognitive load and leakage.

All of the practice transcripts followed a precise script that was analogous to the interview script used in the present experiment. Importantly, raters were trained only to code each detail the first time it appeared within a statement. A detail mentioned in response to two different interview questions was tallied only in response to the first one. The result is that only the amount of novel detail elicited in response to each recall task of the interview was available for assessment. This allows for the tracking of unique details as they are added throughout a statement in order to highlight spontaneous additions. Following the provision of operational definitions, raters coded the details in a practice transcript and then reviewed these codings as a group with the first author.

The raters were given a second practice transcript for homework and reviewed its coding at the second group meeting. This process was repeated for the third transcript. At the final meeting, all discrepancies were resolved and raters were able to have one final round of discussion. Child witness statements from a previous training package (provided to the second author by John Yuille) were used as practice transcripts so that raters would remain blind to the experimental conditions of the present study. Analysis of these statements indicated the following inter-rater reliability for the three detail categories: external = 0.79, contextual = 0.71, and internal = 0.88. The most common mistake was in classifying a detail as contextual versus external, both of which are correlated with honest responding.

JMCQ

The JMCQ is a 36-item questionnaire completed by a trained rater about another person's description of an experienced event. Raters rate each item using a Likert-type scale.

RESULTS

Manipulation checks and outliers

Five interviews had inaudible sections and were excluded from study prior to transcription. Each transcript was read and checked for accuracy according to the interview script, and all interviews followed the RI script. The post-event questionnaire was examined to make certain that all participants responded according to their assigned conditions, and examination of the transcripts verified that all participants did in fact respond in accordance with their assigned condition. There was a statistical issue with sphericity (for Bartlett's test, $p < 0.05$). This was addressed by analyzing the data using Pillai's Trace as a MANOVA.

Anxiety and motivation

The post-event questionnaire contained four Likert-type questions to assess the anxiety and motivation of participants. Both honest and deceptive participants reported a moderate amount of anxiety whilst in the room ($M = 5.2$, $SD = 0.8$). This was comparable to a mean of 5.9 on the question that asked how anxious they are when taking a test in their psychology class ($t(36) = 0.8$; $p > 0.05$). However, deceivers reported being more anxious whilst in room with the exam ($M = 6.3$, $SD = 0.2$) than honest reporters ($M = 4.3$, $SD = 1.4$), $t(36) = 2.6$, $p < 0.05$. They also reported being more anxious during the interview ($M = 5.3$, $SD = 0.6$) than honest reporters ($M = 3.8$, $SD = 0.9$), $t(36) = 2.9$, $p < 0.05$. Overall, participants reported a moderate level of motivation to tell a complete and convincing story. This overall moderate level of motivation masked any potential differences due to the monetary incentive or verbal pep talk, as those in the low motivation condition reported

a similar level of motivation ($M = 4.1$, $SD = 1.4$) as those in the high motivation condition ($M = 4.9$, $SD = 0.8$), $t(36) = 0.6$, $p > 0.05$.

Participant, interviewer, and rater biases

The researchers used discriminant function analysis to classify statements as honest or dishonest on the basis of the amounts of details (external, contextual, and internal) present in and the overall response length of each statement. They then used the canonical variable generated from this analysis to examine any potential biases due to characteristics of the participants. A multiple regression analysis examining the relationship between age, gender, and ethnicity of participants and the canonical variable revealed no significant relationship, $R = 0.15$, $p > 0.05$. A second multiple regression examined potential practice effects in interviewers and raters. Again, no significant relationship was found between the number of interviews administered or transcripts rated and classification accuracy, $R = 0.07$, $p > 0.05$. A third multiple regression examining potential biases in interviewers and raters found no relationship between interviewer or rater identity and classification accuracy, $R = 0.18$, $p > 0.05$.

ACID

ACID yields nine dependent measures for each participant: the response length and the number of external, contextual, and internal details reported during free recall; the response length and the number of new external, contextual, and internal details reported during the mnemonic section of the interview; and whether or not the participant admitted that they could have made a mistake (i.e. four measured at free recall + four measured across mnemonics + one over the entire interview = nine). A multivariate analysis of variance was used to assess for a main effect of the nine ACID criteria across the two reporting conditions (honest versus dishonest). Results indicated a significant effect, Pillai's Trace $(1, 9) = 0.56$, $p < 0.05$, $R = 0.74$. Univariate ANOVAs assessed the individual level of significance of each ACID criterion, with the effect size for each ACID criterion described as a structure coefficient from a discriminant function analysis.

Overall, seven of the nine criteria were significant at the univariate level. Two of the nine criteria did not discriminate significantly between honest and deceptive statements. The means, standard deviations, significance values, and structure coefficients are presented in Table 2.

Finally, the nine ACID criteria were used in a discriminant function analysis to classify statements according to honesty of reporting condition (honest versus deceptive). Overall, 33 of 38 (86.8%) of statements were classified accurately. Fifteen of 19 (78.9%) honest statements were classified accurately, as were 18 of 19 (94.7%) of deceptive statements. To provide a more rigorous estimate of predictive accuracy, cross-validation was performed using the 'leave-one-out' method. In this analysis, 30 of 38 (78.9%) of statements were accurately classified. Thirteen of 19 (68.4%) honest statements were classified accurately, as were 17 of 19 (89.5%) deceptive statements. This second estimate is more likely more accurate, due to the nature of discriminant analysis.

JMCQ

A second multivariate analysis of variance was performed to assess for potential differences in JMCQ scores across the honesty of reporting conditions (honest versus

Table 2.   Means, standard deviations, and effect sizes for ACID criteria (df = 1, 36)

Variable	Honest M (SD)	Dishonest M (SD)	*F*	*p*	SC
Phase/detail type					
Free recall					
External	9.90 (7.50)	11.11 (9.33)	0.19	0.66	0.06
Contextual	4.68 (3.61)	4.11 (3.30)	0.27	0.61	0.08
Internal	1.68 (1.80)	0.32 (0.58)	9.97	<0.01	0.46
Response length	5.03 (2.57)	3.32 (1.71)	5.80	0.02	0.35
Mnemonics					
External	52.05 (21.18)	39.32 (17.86)	4.02	0.04	0.29
Contextual	16.11 (7.96)	11.58 (5.76)	4.03	0.05	0.30
Internal	9.53 (5.12)	3.26 (1.85)	25.10	<0.01	0.73
Response length	20.18 (7.76)	12.37 (6.12)	11.87	<0.01	0.50
Admission of error	1.79 (0.14)	1.30 (0.35)	4.10	0.05	−0.35

ACID, Assessment Criteria Indicative of Deception; SC, structure coefficient.

dishonest). This analysis did not reveal a significant relationship between JMCQ scores and honesty of reporting conditions, $F(12, 25) = 1.93$, $p > 0.05$. At the univariate level, only 3 of the 36 JMCQ questions appeared to differ significantly across the honesty of reporting conditions: Question Number 6, $F(1, 38) = 3.19$, $p < 0.05$; Question Number 12, $F(1, 38) = 5.60$, $p < 0.05$; and Question Number 36, $F(1, 38) = 8.53$, $p < 0.05$. Thus, the relationship between the JMCQ and honesty of reporting appears to be no better than chance.

DISCUSSION

Overall, the system of interviewing and credibility assessment described here was effective at eliciting statements in a manner that facilitated the detection of deception. The accuracy of the system falls somewhere between 79% and 87%. Discriminant functions can inflate accuracy rates by describing and using error variance in predictions, therefore the 79% rate, which is based upon a technique that eliminates the use of individual-specific error variance, is the more accurate of the two. This is significantly better than what would be expected by chance (i.e. 50%). Additionally, untrained raters who read these transcripts achieved a classification rate of 56% (i.e. no better than chance) (Colwell, Hiscock-Anisman, Memon, Colwell, Taylor, & Woods, 2008). The most common strategies of deception were for the student to state that his or her test had fallen from the stack and he or she was being helpful by returning it to the professor's office, or that someone had given him or her permission to go and pick his or her test out of the stack and take it home. Other than these small additions, the liars explained the same series of events as truth-tellers. This is the likely reason that the JMCQ did not uncover differences between groups: All of the participants provided a description based upon a genuine memory.

This study had several additional strengths. First, the experimental task closely mirrored some real-world applications. Second, the participants reported being anxious whilst performing the event. Third, participants were provided with incentive for escaping detection, and they reported a moderate level of overall motivation. Finally, like many real-world

investigations, all participants had time to plan their statement. In these circumstances, the observed 80–90% classification rate is considerable and shows merit for the ACID system.

Investigative interviewing

Because the RI focuses on detecting deception, it creates an interpersonal dynamic that may generate less transfer of control than comparable interviews, such as the Enhanced CI. However, it is likely that the same strategies that facilitate deception detection could hinder the ability of honest reporters, either in the amount or accuracy of information provided. *As such, the authors are proposing the RI only as a means to facilitate the detection of deception, and they do not recommend its use with co-operative witnesses.* Further study of the technique is necessary to determine what effects it has on the amount and accuracy of information obtained from co-operative witnesses. Also, care should be taken when choosing an interview format, and the goal of the interview should guide the nature and structure. However, the RI does facilitate the detection of deception, and it is much less likely to damage memory than some of the confrontational interrogation techniques that often are employed.

Credibility assessment

The content analysis portion of the ACID system performed quite well overall. Seven of nine traditional ACID variables were related significantly to honesty or deception. Deceptive statements were shorter, less detailed, and had fewer details provided in response to the mnemonic portions of the RI, whereas honest reports were longer, more detailed, and had more details provided in response to the mnemonic portions of the RI. The largest differences were seen when comparing the mnemonic portions of the interviews: The mnemonic section of honest statements consistently contained more details and was longer than the mnemonic section of deceptive statements. This pattern was consistent with observed differences in vividness and spontaneity described in previous research (Colwell, Hiscock-Anisman, Memon, *et al.*, 2007). Honest reporters also were more likely than deceivers to admit possible error (two-thirds of honest reporters versus one-third of deceivers). As a whole, the differences seen here are consistent with explanations based upon memory quality and the desire to control information (Porter & Yuille, 1996; Vrij & Mann, 2006).

   The mnemonics portion of the RI enhances recall on behalf of honest respondents and, therefore, leads to relatively longer responses with more new details. Conversely, deceivers often give responses based upon their careful lie script and do not try to answer based upon the real event; thus, they do not benefit from such retrieval cues. In the current study, the mnemonic portion of the interview for deceivers was relatively shorter and contained relatively fewer additional details. Thus, it appears that the improvement in recall provided by mnemonics represents a difficult task that actually diminishes the performance of deceivers (Colwell *et al.*, 2002; Colwell, Hiscock-Anisman, Memon, *et al.*, 2007; Porter & Yuille, 1996).

   In other words, when deceivers were faced with a potentially unrehearsed lying task and a potentially higher cognitive load, they provided shorter responses with much less additional external or contextual detail when compared to honest respondents. Deceivers tended to provide the same basic lie script in response to each question, whereas the statements of honest reporters grew and evolved as a result of the mnemonics. This is

consistent with the repeat versus reconstruct hypothesis (Granhag *et al.*, 2003; Granhag & Strömwall, 1999). This makes the structure of the interview very important. The free recall portion of the RI served as a baseline to indicate a participant's schema or script for the event. Deceivers did not elaborate upon this script due to the mnemonics, whereas honest reporters did. Therefore, administration of the free recall task prior to administration of the mnemonics tasks may be an effective strategy for the detection of deception.

In this study, inferential questions were included as part of the RI, due to their ability to increase cognitive load and to hinder superficial encoding, thereby facilitating the detection of deception (Colwell *et al.*, 2002). At present, it is uncertain what effect, if any, these questions have on honest respondents. To be safe, these inferential questions should not be introduced until after the mental reinstatement of context mnemonic to avoid the possible addition of unknown effects. In practice, these questions also can be used to form a forced-choice guilty knowledge task, which should increase further respondents' anxiety as well as provide another indicator of deception or credibility. Still, research is necessary to determine whether these questions have a negative impact on the accuracy or amount of information provided by honest respondents. At present they are recommended only when the goal of an interview is the detection of deception.

Comparing ACID to JMCQ

Both the content criteria of the ACID technique and the JMCQ were derived from RM research. Therefore, it seems somewhat contradictory that one should outperform the other. It may be that the JMCQ performed relatively poorly in the current study because all of the statements essentially described an experienced event. The JMCQ may work better when attempting to distinguish imagination or fabrication from honest responding (Sporer, 1997). Also, although initially derived from RM, the content criteria used in ACID are operationalised in a manner that allows for an assessment of the process of deception. The JMCQ focuses mostly on overall statement characteristics and, therefore, misses information regarding how honest statements change with additional recall tasks. Finally, the content criteria used in the ACID system are operationalised as ratio-level data, and this may result in greater statistical power and better sensitivity and specificity of classification.

Contrary findings

Contrary to our expectations, honest reports contained *more* internal details than deceptive reports. This goes against the supposition of early RM research but was likely due to the fact that emotional details were counted in the internal category. Early RM hypothesised that confabulations would have more internal or idiosyncratic information because they were dealing with alcoholic delusions or memory errors made by witnesses. In these cases, the mistakes in RM were likely to arise when the details in question were similar to the witness' previous experience. Similarly, IMT postulates that deceivers would provide information from past memories (which are rated as internal and idiosyncratic) in order to give the appearance of co-operation without actually disclosing any potentially damaging information from the time of the theft. In our study, however, a large proportion of internal details dealt with emotional content, particularly anxiety experienced whilst in the office. As reported earlier, deceivers felt more anxiety during the interview, but they were trying to report in such a way that concealed their crime. In this situation, reporting anxiety

could provide an indicator to the interviewer or rater that the participant was engaged in some wrongdoing. Because they did not want to give indication that they were doing anything wrong, they should have avoided mentioning any experience of anxiety during the event. Honest reporters, who had nothing to hide, were more willing to report that they were anxious. Therefore, it may be that honest reporters are more likely to report emotional detail overall, or it may be an artefact of the present study. Despite being contrary to predictions, this relationship between non-disclosure of emotion and deception is consistent with previous research (Colwell, Hiscock-Anisman, Memon *et al.*, 2007). Future research should examine more carefully the relationship between participants' descriptions of their own emotional state and honest responding. If the finding bears out, future versions of the ACID technique may need to separate emotion or affect from other internal details.

Conclusions

This research supported the combination of investigative interviewing and credibility assessment described as the ACID technique. This combines the RI, which increases cognitive load and highlights attempted control of information, with content criteria derived from memory and impression management research. ACID accurately classified the majority of statements as honest or deceptive. This is consistent with other research regarding all or parts of this system when used with inmates, college students, or children (Colwell *et al.*, 2002; Colwell, Hiscock-Anisman, Leach *et al.*, 2007; Colwell, Hiscock-Anisman, Memon *et al.*, 2007; Memon *et al.*, 2008). In the present study, all participants engaged in the same tasks, so the only differences in the statements resulted from the attempts at deception. These differences were so subtle that untrained raters could not detect them, nor could the JMCQ.

The major limitation of the current research was the inclusion of affective details in the same category as internal details. Because of this, the operational definitions of the ACID system have been modified, and future iterations will have a separate category for affective details (Memon *et al.*, 2008).

This simulated theft scenario had several characteristics consistent with a true investigation, including anxiety whilst performing the task, motivation not to be seen whilst committing the act, and motivation on behalf of respondents to provide a story that convinced the interviewer. Unfortunately, as with most credibility assessment studies, the intermediate level of motivation present in this study is different from the high levels that often exist in forensic settings. However, previous research (DePaulo & Kirkendol, 1989; Vrij, 2006) suggests that, if anything, the increased motivation that comes with real-world, high-stakes lie situations, as in criminal investigations, should enhance detecting deception.

As always, additional research will be needed to further validate the ACID technique. Future studies should address optimal interviewing, other levels of motivation to report or to escape detection, and different applications and populations. The last two are possibly the most promising. ACID is currently being studied as an interviewing and assessment tool to increase the ability of professionals to detect deception when reading transcribed investigative interviews (Colwell *et al.*, 2008), and it is also being used as an interviewing and assessment tool in languages other than English. Thus, the ACID system represents a promising new tool in investigative psychology, one that combines theory and research in investigative interviewing, memory, and impression management to facilitate the detection of deception in suspect and witness statements.

REFERENCES

Colwell, K., Hiscock-Anisman, C. K., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology*, *16*, 287–300.

Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. M. (2006). Strategies of impression management among deceivers and truth-tellers: How liars attempt to convince. *American Journal of Forensic Psychology*, *24*, 31–38.

Colwell, K., Hiscock-Anisman, C., Leach, A., Corbett, L., & Uerz, E. (2007). Reality Monitoring criteria and real versus imagined autobiographical memory in children. Manuscript in preparation.

Colwell, K., Hiscock-Anisman, C., Memon, A., Rachel, A., & Colwell, L. (2007). Vividness and spontaneity of statement detail characteristics as predictors of witness credibility. *American Journal of Forensic Psychology*, *25*, 5–30.

Colwell, K., Hiscock-Anisman, C., Memon, A., Colwell, L. H., Taylor, L., & Woods, D. (2008). Training in Assessment Criteria Indicative of Deception (ACID) to improve credibility assessment. Manuscript under review.

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51–70). New York: Kluwer Academic/Plenum Publishers.

De Paulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

Fisher, R., & Geiselman, R. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL, England: Charles C. Thomas.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology*, *70*, 401–412.

Granhag, P., & Strömwall, L. (1999). Repeated interrogations: Stretching the deception detection paradigm. *Expert Evidence*, *7*, 163–174.

Granhag, P., Strömwall, L., & Jonsson, A. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology*, *33*, 848–868.

Hiscock-Anisman, C., Colwell, K., Memon, A., & Laurita, A. (2008). Using the Reality Interview to improve credibility assessment. Manuscript under review.

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

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371–376.

Mann, S. M., Vrij, A., & Bull, R. (2002). Suspects, lies, and videotape: An analysis of authentic high-stake liars. *Law and Human Behavior*, *26*, 365–376.

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

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, *59*, 1–16.

Memon, A., Mastroberardino, S., Fraser, J., & Colwell, K. (2008). Reality Monitoring and credibility assessment. Manuscript under review.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, *20*, 443–458.

Sporer, S. L. (1997). The less traveled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, *11*, 373–397.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. West Sussex, England: John Wiley & Sons.

Vrij, A. (2006). Challenging interviewees during interviews: The potential effects on lie detection. *Psychology, Crime, & Law*, *12*, 193–206.

Vrij, A., & Mann, S. M. (2006). Criteria-Based, Content Analysis: An empirical test of its underlying processes. *Psychology, Crime, & Law*, *12*, 337–349.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, *10*, 141–142.