Chapter 11
Assessment Criteria Indicative of Deception: An Example of the New Paradigm of Differential Recall Enhancement

Kevin Colwell, Cheryl Hiscock-Anisman, and Jacquelyn Fede

The present chapter details the historical and conceptual evolution of a new paradigm in statement analysis that has developed over the past 20 years. There has been an increasing awareness of the importance of interviewing designed to facilitate the detection of deception as a necessary component of statement analysis (Colwell, Hiscock, & Memon, 2002; Hartwig & Bond, 2011; Hernández-Fernaud & Alonso-Quecuty, 1997; Koehnken, Schimossek, Ascherman, & Hofer, 1995; Vrij, Fisher, Mann, & Leal, 2006). Subsequently, the work of multiple researchers has created a zeitgeist that has nurtured and informed the development of this new paradigm. This chapter begins by providing a quick overview of the various lines of research that comprise this paradigm. Attention is then given to credibility assessment and statement content criteria that discriminate honest from deceptive responding. Then, the focus is on strategies of impression management and the subjective experience of respondents during an investigative interview. This sets the stage for a discussion of investigative interviewing structure and techniques that facilitate the detection of deception through the process of Differential Recall Enhancement (DRE: Colwell et al., 2012). Finally, this chapter considers in detail an approach to interviewing and assessment that is representative of the new paradigm.

The New Paradigm

In the past two decades, there has been a shift to focus on the importance of interviewing to detect deception as being the most important aspect of statement analysis (Hartwig & Bond, 2011). Without effective interviewing, there are few reliable

K. Colwell (✉) • J. Fede
Department of Psychology, Southern Connecticut State University,
501 Crescent St., New Haven, CT, 06515 USA
e-mail: colwellk2@southernct.edu

C. Hiscock-Anisman
The National University,
La Jolla, CA, 92037 USA

B.S. Cooper et al. (eds.), *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment*, DOI 10.1007/978-1-4614-5547-9_11,

differences between honest and deceptive responding. With effective interviewing and an awareness of critical content criteria, the differences between honest and deceptive responding are maximized and readily apparent. This work can be traced back to independent studies done comparing the Step-Wise Interview (SI: Zaparniuk, Yuille, & Taylor, 1995) to the Cognitive Interview (CI: Fisher, Geiselman, & Amador, 1989) or the SI, CI, and Reality Interview (Colwell, 1997). The first of these studies noted that adult memory for events is so complex, in general, that effective credibility assessment will require the use of interviewing that enhances differences between honest and deceptive responding (Koehnken et al., 1995). The second went on to state that the relationship between question type and content criteria should be closely studied (Colwell, 1997), with the intent of using techniques that take advantage of the increased cognitive and interpersonal demands placed upon deceivers (Colwell et al., 2002). In recent years, there has been a proliferation of research based upon these ideas. There has been the *cognitive load* hypothesis (Vrij et al., 2006), which builds upon the early work by focusing upon the specific CI technique of reverse-order recall to magnify differences between honest and deceptive responding. Similarly, unanticipated questions have been shown to increase cognitive load (Vrij et al., 2009). In the same vein, there has been study of the manner and timing of disclosure of evidence during an investigation to facilitate the detection of deception (Dando & Bull, 2011; Hartwig, Granhag, Strömwall, & Kronkvist, 2006). The oldest and most integrated set of techniques that represent this paradigm is assessment criteria indicative of deception (ACID: Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2008). All of these approaches share a goal of enhancing recall for honest respondents while making deception more difficult for deceivers, and thereby magnifying the differences between the two. This is the essence of DRE – to help honest respondents while making deception more difficult and more obvious.

Credibility Assessment

The process of statement analysis refers to the use of content criteria in the analysis of a statement taken from an investigative interview. This process involves a properly conducted interview, content analysis of the resulting statement, and careful analysis of all available case data. This is drastically different from interrogation, which is neither ethically permissible for psychologists in the United States nor designed as an investigative tool. Statement analysis, first and foremost, seeks to obtain accurate and useful information from victims, witnesses, and suspects. In other words, it is primarily an investigative tool. Secondarily, statement analysis seeks to provide a mechanism for assessing the credibility of the information obtained. Credibility assessment determines whether a statement possesses the characteristics associated with accurate recall for an experienced event. It is related to detecting deception, but there are some specific differences. Credibility assessment is a form of truth confirmation. It seeks to provide a mechanism for weighing the various sources of information presented to an investigator or to a trier of fact.

Especially in the legal systems derived from British Common Law (e.g., Australia, US, Canada, United Kingdom), whether a statement is honest or deceptive is a decision for the trier of fact (e.g., judge, jury, magistrate; see Seniuk, present volume). It also provides a mechanism for investigators to determine what additional information must be gathered. This last application is the primary role for the type of credibility assessment described in the present chapter.

Memory and Credibility Assessment

Criteria-Based Content Analysis

The oldest, most researched, and prototypical approach to statement analysis is Criteria-Based Content Analysis (CBCA; Vrij, 2005). This system was first devised for use with allegations of child sexual abuse in Germany, and it has been used as part of court-mandated assessments since the 1950s (Undeutsch, 1954). The underlying premise of CBCA is that systematic differences exist between statements derived from memory for a real event and statements derived from imagination or fabrication. This has been referred to as the *Undeutsch hypothesis* (Porter & Yuille, 1995; also see O'Sullivan, present volume). There has been debate in the field as to the total number and application of CBCA criteria. However, it is generally accepted that CBCA comprises at least 19 content criteria, the presence of which increases the likelihood that a statement is true (Zaparniuk et al., 1995). CBCA is a form of credibility assessment – higher numbers of the criteria do not indicate honesty per se but rather increase the likelihood that the statement is derived from genuine experience, that is, it is more likely to be honest (see Griesel, Ternes, Schraml, Cooper, & Yuille; present volume). Under certain circumstances, a single criterion could suffice for the statement to be deemed credible.

CBCA was designed to evaluate statements from children regarding alleged abuse. For this reason, a number of the criteria are not relevant to all statements. Because the present work focuses on interviewing and credibility assessment in general, we limit our discussion to the portion of CBCA that applies to all memories for events, not just to alleged victims' memories for child abuse.

The presence of the first three criteria of CBCA is considered to be necessary in order for a statement to be judged as credible, and these are the three that apply to memories for all events. These criteria include: coherence (sometimes referred to as logical structure), sufficient detail (sometimes referred to as appropriate amount of detail), and spontaneous reproduction (sometimes referred to as unstructured production; Zaparniuk et al., 1995). Coherence (or logical structure) refers to the various portions of a statement that consistently hold together and agree with one another. It also deals with whether the events described in a statement are possible given the basic limitations of time and space. Therefore, a statement that contains serious contradictions, or one that is simply physically impossible, would not be

rated as coherent. In contrast, a statement in which all of the portions describe the same basic event in the same basic manner, and which restricts itself to limitations imposed by time and space, would be rated as coherent (Colwell et al., 2002). This criterion is especially important in the assessment of children's statements. Children often accidentally release sensitive information which contradicts information they have previously stated (Williams et al., 2012). Coherence has been found to have some utility with adults (Colwell et al., 2002). However, this criterion should be applied carefully to adult statements. For example, one study found that honest men are more likely to provide incoherent statements than are deceptive women (Suckle-Nelson et al., 2010).

Sufficient detail is a statement characteristic that addresses the rich amount of sensory information that can be provided by a cooperative witness who is reporting an event he or she experienced. Credible statements tend to contain a copious amount of detail and are rich in visual, spatial, and auditory information (Vrij, 2005). Determining the amount of detail that is sufficient is a subjective judgment made by the rater, and it is based upon experience and training. There are issues with training, reliability, and confounds due to age and language (Blandon-Gitlin, Pezdek, Lindsay, & Hagen, 2009; Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005). Therefore, this criterion assesses a vital aspect of credibility, but there are problems with its current application.

One major limitation of the sufficient detail criterion comes from the manner in which CBCA is scored. Therefore, this limitation will apply to all CBCA criteria, but is considered here. In CBCA scoring, the criteria are scored as *present* or *absent*, and they are scored so that, if they appear anywhere in a statement, they are counted as, "present" for that statement. This type of scoring loses the rich information that can be gained by matching the type of question asked with resulting content criteria (Colwell et al., 2002). This dichotomous scoring reduces the extent to which variability is possible and negatively affects the psychometric properties of the criteria; that is, it reduces the reliability and potential validity. The scoring of a statement as a whole minimizes the role of the all-important effects of interviewing and the requisite understanding of how memory operates. That is, different questions lead to different statement characteristics, and it is important to link questioning strategy to content criteria to achieve optimal results (Colwell et al.).

Spontaneous reproduction (or unstructured production) is a statement characteristic that addresses the offhand and unplanned nature of honest responding (Zaparniuk et al., 1995). This characteristic emerges for two reasons. First, honest respondents are aware of their honesty and might believe that other people can see this sense of honesty. So, they are not concerned with telling a scripted narrative from start to finish, and are free to provide information as they remember it (Colwell, Hiscock-Anisman, Memon, Yaeger, & Michlik, 2006; Hines et al., 2010). The second reason has to do with the automatic nature of memory. The process of interviewing, when carried out correctly, leads to the recall of additional information (see Yarbrough, Hervé, & Harms, present volume). Therefore, a person who is engaged in honest responding will provide a certain amount of detail in response to an initial question. If additional questions are asked, especially questions using

mnemonics, then the honest respondent should have a significant amount of additional information become available due to spreading activation and cue-dependent recall (Colwell et al., 2008; Fisher et al., 1989; Memon, Fraser, Colwell, Odino, & Mastroberadino, 2009; Memon, Meissner, & Fraser, 2010). In short, honest responding during an investigative interview leads to a positive-feedback system. The act of remembering provides new cues, which in turn lead to even more remembering.

Table 11.1 contains an overview of CBCA studies and their findings regarding sufficient detail and spontaneous reproduction, as these are the two criteria that appear to be the most promising indicators of credibility across a range of ages and with both genders. Coherence is not considered in this table due to potential misapplication with adults (Suckle-Nelson et al., 2010).

Although there has been some promise in CBCA research, there are many problems that hinder its application in North America. For example, CBCA is confounded by age, familiarity with the type of event, and language capacity (Blandon-Gitlin et al., 2009, 2005; Vrij, Akehurst, Soukara, & Bull, 2002). Therefore, the system of adult credibility assessment presented here pulled from the theory underlying CBCA but was forced to consider other perspectives related to memory in order to avoid the pitfalls described above (dichotomous scoring, confounds with age, gender, and, uncertain rules for when a statement should be labeled, "credible," or, "not credible").

Reality Monitoring

Johnson and Raye (1981) posited that memories for experienced events will have more external-sensorial information and more contextual information than will memories derived from imagination or fabrication. The method of assessment based on that hypothesis, labeled Reality Monitoring (RM), initially appeared to be promising. Indeed, numerous studies found a direct relationship between the amount of sensory detail and the credibility of a statement (Masip, Sporer, Garrido, & Herrero, 2005). RM assessments have been used to assess the credibility of intrapersonal memories and interpersonal statements. In the first case, a person assesses his or her own memory and, in the second, an external rater typically reads a transcript from a statement and rates it. The ratings have been done according to Likert-type scales or by tallying the amount of individual details related to sensorial, contextual, and internal cognitive processes. These Likert-type and the detail tally assessments, despite apparent differences, are actually assessing the same constructs and perform with the same level of accuracy (Memon et al., 2009).

In general, RM-based techniques have led to accuracy rates in the 80% range when predicting statements as honest or deceptive (Masip et al., 2005). The original hypothesis of RM, when applied to interpersonal deception, also posited that deceptive statements will have more details derived from internal sources, such as cognitive operations, imagination, fabrication, associated memories for previous events, etc. Unfortunately, this hypothesis has not been supported (Memon et al., 2009;

Table 11.1 Studies considering the relationship of sufficient detail and spontaneous reproduction to credibility

Citation	Sufficient detail	Spontaneous reproduction
Akehurst, Koehnken, and Hofer (1995)	Increase	No relationship
Blandon-Gitlin et al. (2009)		
Experiment 1	Increase	No relationship
Experiment 2	No relationship	Increase
Boychuk (1991)	Increase	Increase
Esplin, Boychuk, and Raskin (1988)	Increase	Increase
Hofer, Akehurst, and Metzger (1996)	Increase	No relationship
Koehnken et al. (1995)	Increase	Increase
Lamb, Sternberg, Esplin, and Hershkowitz (1997)	Increase	Increase
Landry and Brigham (1992)	Increase	*
Porter and Yuille (1996)	Increase	No relationship
Ruby and Brigham (1998)	No relationship	Decrease
Steller, Wellershaus, and Wolf (1988)	Increase	*
Vrij et al. (2002)	Increase	*
Vrij, Edward, Roberts, and Bull (2000)		
Experiment 1	No relationship	*
Experiment 2	Increase	No relationship
Winkel and Vrij (1995)	Increase	Increase
Zaparniuk et al. (1995)	No relationship	Increase
Totals	13 increase	7 increase
	4 no relationship	4 no relationship: 1 decrease

Note: An asterisk (*) indicates that no data was available for this study

Memon, Omerod, & Dando, 2012). Furthermore, there are a number of problems with the measurement of RM that hinder its forensic application (e.g., lack of consistent definitions, poor reliability, no accepted decision criteria, confoundedness with the emotional valence of an event; Memon et al., 2012).

What, then, was the reason for the initial success of RM-based assessments? RM appeared to be promising because, during an investigative interview, honest respondents often provide more overall detail than do deceptive respondents. This increase in detail allows for classification at higher-than-chance rates. Because the overall amount of detail in a statement, (especially under appropriate interviewing circumstances) is correlated with the credibility of a statement, it often appeared that the assessment of RM could provide an effective mechanism of credibility assessment (Colwell, Hiscock-Anisman, Memon, Rachel, & Colwell, 2007). However, the increase in detail was not due specifically to the reasons posited by RM theory. Honest statements do have more sensory details in many circumstances, but deceptive statements do not have more details from cognitive operations, previous memories, or other internal sources (Colwell et al., 2007). Finally, and to restate the central lesson of this chapter, no content criterion ought to be considered in the absence of the interviewing technique used to elicit the statement (Colwell et al., 2002).

Honest respondents may provide more detail, or they may provide less, depending upon the question posed (Memon et al., 2009).

Impression Management and Credibility Assessment: Subjective Clues

Credibility assessment must take into account subjective indicators of deception (i.e., those behaviors that people believe to be indicative of honesty or deception) as well as objective indicators of deception (i.e., those behaviors that truly are indicative of honesty or deception). This is due to the central fact that deceivers are aware of their deception and take steps to hide this. Deception during an investigative interview is a pragmatic enterprise where deceivers must avoid the disclosure of sensitive information, avoid making obvious contradictions in their statements, and generally appear honest and cooperative. They must present sufficient information to satisfy their interviewer while withholding or changing any information that could lead to their detection (Colwell & Sjerven, 2005; Hines et al., 2010; Porter & Yuille, 1996). Therefore, a complete approach to interviewing and credibility assessment needs to account for the effort that is being made by the respondent to avoid detection and to look honest. Knowing what people think is indicative of deception, and where people are exerting effort provides important information for crafting a system of interviewing and assessment. Most important are discrepancies between what is thought to be indicative of credibility and what truly is indicative of credibility. This allows the assessor to judge credibility without worrying about the effects of motivation or preparation. Often, in fact, motivation and preparation can make deception detection easier because the effort expended by motivated deceivers leads to predictable changes in their behavior, whereas motivation does not have the same effects on honest respondents (Colwell et al., 2002, 2007).

One of the best ways to determine how honest and deceptive respondents attempt to present honestly during an investigative interview is to ask them. In a series of studies, Hiscock-Anisman et al. (2012) did exactly this. College students either committed or witnessed a theft, or either told the truth or lied about an autobiographical memory, and then underwent an investigative interview. Several hundred students from universities across the US have been assessed. The demographic and socio-economic characteristics of these samples are quite varied, including one set of students who spoke Chinese and had their data translated into English for assessment. Table 11.2 provides a summary of the strategy data provided across all of these different studies. These studies generally involved the chance to win up to $200.00 for successfully convincing an interviewer that one was honest. This level of motivation is consistent with a large number of situations, but it does not match the extreme consequences of some investigative situations. However, the vast literature on the relationship between performance and anxiety is clear on the point that people do not develop and demonstrate new skills when under high

stress. They simply continue to use the strategies they have already learned, with a decrease in skill level as the anxiety or consequences move from moderate to high. Students are motivated at a moderate level and, in fact, it is difficult to find non-motivated student volunteers. Therefore, the information obtained from the assessment of students, despite the doubts of many investigators, provides a very good insight into honesty and deception during investigative interviews (Colwell et al., 2002).

One important finding from the above data is that deceivers tried hard not to make mistakes in their stories, and they incorporated some specific strategies in doing so. Deceivers were generally concerned about making certain that they did not do anything that would draw attention to their story, such as to present any inconsistencies or to make overt mistakes. They also developed and practiced their stories in advance of the interview. They believed that this is a useful approach, to 'stay on script', in order not to provide any information that might implicate them. This approach also allowed the deceivers to feel better prepared to answer questions based upon the fabricated script rather than upon the real event in question. Deceivers wanted to make sure that they presented both verbal and nonverbal information in a controlled way. They believed that, if they manage the information, they would be less likely viewed with suspicion. Deceivers also believed that it is important to be seen as cooperative as possible and to avoid drawing attention to themselves. This was done by appearing calm and sincere and by acting certain about the information presented. Strategies such as appearing relaxed, appearing self-assured, and providing direct eye contact were viewed as important approaches to avoid being caught. Overall, deceivers in this study wanted to provide a relatively short, carefully phrased description and to appear confident while doing so. Many strategies which were listed by deceptive respondents were also listed by honest respondents. In other words, there was considerable overlap between the intended behavior of honest and deceptive respondents during an investigative interview. However, it is possible to elicit differences between the two groups through careful, strategic interviewing. The data from this large, multi-site and multi-ethnic sample is also consistent with two previous studies on this same topic (Colwell, Hiscock-Anisman, et al., 2006; Hines et al., 2010).

Perhaps the most important finding from research regarding subjective strategies of deception is the mismatch between perceived and genuine cues to credibility or deception (see ten Brinke & Porter, present volume). The following information comes from Table 11.2 as well as from a previous series of studies (Colwell, Hiscock-Anisman, et al., 2006; Hines et al., 2010). More than 75% of respondents have a wrong understanding of the relationship between the amount of detail in a statement and the credibility of that statement. Only one participant across all samples correctly mentioned that adding information as the interview progressed was indicative of credibility. In contrast to these perceptions, the most powerful predictor of honest responding in our research, given appropriate interviewing, is the addition of new details following an initial description. Similarly, deceptive statements

Table 11.2 Strategies of impression management for honest and deceptive participants ($N = 320$)

Strategy to appear credible	Honest % ($n = 175$)	Deceptive % ($n = 145$)	t	p
Details not mentioned	38	49	0.26	0.79
Details mentioned, direction not specified[a]	15	15	0.03	0.98
Complete detail	32	26	0.91	0.36
Minimal detail	9	10	−0.22	0.82
Calm and confident	28	35	−1.20	0.24
Coherent and consistent	8	41	−1.60	0.01
Thoughts and emotions[b]	17	9	1.80	0.08
Eye contact	7	17	−2.30	0.02
Accuracy of details provided	11	52	1.20	0.01
Honest about non-event details	11	14	0.18	0.50
Tone of voice	4	3	0.58	0.78
Convincing or plausible	2	0	1.60	0.12
Believe it yourself	4	3	0.64	0.52
Spontaneous (credible)	1	0	1.10	0.27
Not Spontaneous	3	3	−0.11	0.91
Other	4	4	−0.05	0.96

[a]These participants mentioned that statements should have detail but they did not say whether a high or a low degree of detail gave the appearance of credibility

[b]These participants indicated that one should describe either what one was thinking or what one was feeling during the target event in order to appear credible

become obvious because they are shorter and more carefully phrased. This means that, to the extent that participants have insight into their own behavior, they are focusing their efforts in the wrong areas. This mismatch between perceived and genuine cues, and the resulting misplaced effort, minimizes the benefits of planning and motivation. In fact, it is likely that planning and motivation have a paradoxical effect, leading to increased ability to detect deception through appropriate interviewing and assessment. A similar situation has long been observed in the symptom validity approach to the assessment of malingering (Colwell & Sjerven, 2005; Hiscock & Hiscock, 1989).[1]

[1] In the symptom-validity approach to malingering, respondents who are motivated often perform worse-than-chance on two-alternative, forced-choice tests.

Impression Management and Credibility Assessment: Objective Clues

The Lie Script

One primary goal of impression management during an investigative interview is to provide a story that does not include any information that could lead to detection (for deceivers) and to provide a narrative that does not have any major contradictions (for deceivers and honest respondents, alike). Most honest respondents tend to believe that their honesty is transparent, whereas most deceptive respondents believe that they must plan ahead and prepare in order to appear honest (Hartwig & Doering, 2009). This preparation is often the development of a fictitious account of the target event. This account, rather than a memory for a real event, is used to provide information to investigators. In this way, a deceiver can avoid sensitive information, can give the appearance of cooperation, and can provide a story that is coherent and well phrased. This strategy was originally termed *superficial encoding* (Porter & Yuille, 1995), and it is currently referred to in the literature as the use of a *lie script* (Colwell, Hiscock-Anisman, et al., 2006). Regardless of the term used to describe it, this strategy has two components: the creation and rehearsal of a lie script to replace the target memory, and the inhibition of the target memory during the interview (note: even when people are lying to claim that they have done things they did not, they have a memory for what they really did during that time period, and they must inhibit this memory and replace it with their lie script). In regard to the first component, some research has demonstrated that deceptive responses during an investigative interview are often shorter, are more carefully phrased, and contain less unique detail than honest responses (Colwell et al., 2007; Suckle-Nelson et al., 2010). In regard to the second component, some brain-imaging studies suggest that there is a significant amount of activity in inhibitory cortical centers during the act of deception, suggesting that many deceivers expend mental effort to inhibit their memory for the target event (Karim et al., 2010). This need to inhibit the memory for the original event for successful deception has important implications for investigative interviewing and is discussed at length in that section below.

The use of a lie script to avoid detection (e.g., due to accidental disclosure of sensitive information or to contradictions in one's story) seems to develop between childhood and adolescence. These were among the most common mistakes that children made when attempting deception in previous research. Williams et al. (2012) found that between 20% and 30% of children, between the ages of 8 and 12, who were lying about taking a wallet during a scavenger hunt either accidentally disclosed sensitive information or gave a story with major contradictions. These children typically either mentioned the wallet when they should not have, or changed their story midway through the interview. Such deceptions are obvious. As adults, we have learned this lesson well and, therefore, focus a great deal of our attention on avoiding the release of sensitive information or on not contradicting ourselves (Suckle-Nelson et al., 2010). This often is taken so far that many deceivers believe

that making changes to their story or admitting that they could be mistaken is actually indicative of deception. Again, the use of a short script helps the deceiver with his or her task. Also, providing essentially the same script to each question asked is common (Colwell, Hiscock-Anisman, et al., 2006; Hines et al., 2010).

There is one significant result of using a lie script that has implications for investigative interviewing and credibility assessment. The act of rehearsing a script and then answering questions based upon that script (rather than upon the original memory) results in a loss or change of information in the original memory. In other words, the act of deception may change one's memory for an event (Colwell, Hiscock-Anisman, Corbett, et al., 2011). People who are lying to say that they did something that they did not actually do may come to believe that they did this thing, while those lying to omit an action that they actually committed may come to believe that they did not do this thing. The implications of this may be profound. The act of holding suspects for long-term interrogation may not only be unethical but also fruitless. Rehearsing a deception appears to be akin to imagination inflation (Loftus & Palmer, 1974) and may preclude later access to accurate information. Therefore, a person who spends a period of time carefully rehearsing a lie to fool interrogators may never be able to remember accurately the true information that is sought by those interrogators. Not only will many suspects lie to escape captivity, but those who may eventually desire to tell the truth are likely unable to provide accurate information after their period of internment. Additionally, holding innocent people for long periods of time, especially in conditions that promote anxiety, could lead to continued rehearsal. This continued rehearsal of their honest statements could make these honest statements become more rigid and cause them to appear more like deceptive statements. Long-term confinement can mask differences between honesty and deception and may render a person relatively useless as a potential source of information.

Appearing Calm and Cooperative

The secondary goal of impression management, on behalf of deceivers, is arguably to attempt to appear calm and cooperative. A well-spoken, confident response is considered to be a clue to honesty (Colwell, Hiscock-Anisman, et al., 2006; Hines et al., 2010). To the untrained observer, a short and carefully phrased script facilitates this sort of responding. Having a script that excludes any information that could lead to detection allows the deceptive respondent to be less anxious than he or she would be if forced to create the lie during the interview. In addition, a clear and well-organized response conveys the impression of credibility and certainty. The deceptive respondent is then able to appear confident. In previous research, this confidence has been described in two related manners – one deals with lack of anxiety, as in, "calm and confident," and the other deals with a metacognitive assessment, as in, "certain about the correctness of their statement." It is arguably vital to convey both variations of confidence if one is to appear credible according

to the average person (as well as to law enforcement; Colwell, Miller, Miller, & Lyons, 2006).

Careful phrasing can be measured by the type-token ratio (TTR), which is a ratio of the unique words in a statement to the total number of words in a statement. For example, the sentence "One small step for man, one giant leap for mankind," has a TTR of 0.8. There are eight unique words in the statement and ten total words. As people speak more carefully, they tend to speak with a wider range of their vocabulary in order to look intelligent and helpful, and they tend to provide fewer total words to avoid the possibility of making a mistake. Both of these tendencies cause the TTR of deceptive respondents to be higher than the TTR of honest respondents during an investigative interview (Colwell et al., 2002). Honest statements tend to be long and not-so-careful, whereas deceptive statements tend to be short and careful. Along the same lines, deceivers are less likely to admit that they could have been mistaken than are honest respondents. This finding is also consistent with the motivational criteria from CBCA, most specifically, "admitting lack of memory" (Griesel et al., present volume; Zaparniuk et al., 1995). Unfortunately, like any other cue, willingness to admit error is not diagnostic by itself. Approximately one sixth of deceivers will admit they could have been mistaken, whereas approximately one third of honest respondents will admit to such potential error (Colwell et al., 2008). The application of this criterion parallels the larger state of affairs in investigative interviewing and credibility assessment. No single criterion is indicative of honesty or deception, and there must always be a careful consideration given to (1) other aspects of a statement, and (2) all other available case data. A significant amount of hardship and mistaken decision making could have been avoided if investigators had always realized that a single criterion (e.g., looking up and to the right or to the left) is not indicative of honesty or deception (Scheck, Neufeld, & Dwyer, 2001).

Summary

In sum, research has shown that most deceptive statements tend to be shorter, to be less detailed, to be more carefully phrased, to contain more contradictions or sensitive disclosures (especially in children), and to be less likely to contain admissions of possible mistakes. Research targeting the subjective perceptions of those engaged in an investigative interview has shown that honest respondents believe that their honesty should be transparent; therefore, they do not expend as much effort in managing their appearance. In contrast, deceivers work to manage their appearance by creating and rehearsing a short script to avoid incrimination and to appear cooperative. While responding, deceivers attempt to avoid sensitive disclosures, contradictions, or anything that would create questions regarding their credibility, such as changing their story, admitting mistakes, or appearing anxious. Taken together, this indicates that the act of deception is a more difficult and planned act than that of honest responding. Honest respondents are free to access their memory for the original event, whereas deceivers must constantly control information and attempt to

stick to their rehearsed lie script. Therefore, honest respondents have fewer demands placed upon them and are able to benefit from the positive-feedback nature of recall, whereas deceivers are under significant cognitive demands and are not able to benefit from the positive-feedback nature of recall.

It is now possible to provide an integrated approach to content criteria and credibility. Often, honest statements during an investigative interview are more detailed, and they tend to have more words and more unique details added after the first telling of the story. The fact that there is more overall detail relates in part to vividness, and the fact that more words and more unique details are added after the free recall relates in part to spontaneity (Colwell et al., 2007). This is because honest recall is an automatic process that forms a positive-feedback mechanism. The act of remembering leads to the recall of new information, which can then be used as recall cues for even more information. The result is more and more information from honest respondents as an interview progresses. In contrast, deceptive statements are often less detailed, and they have significantly fewer words and details added after their initial free recall. This is because most deceivers believe that adding new words and details after free recall causes suspicion in interviewers (Colwell, Hiscock-Anisman, et al., 2006; Hines et al., 2010). Another reason that fewer new words and details are added by deceivers after an initial free recall is that deceivers are focusing their effort on providing careful phrasing and on tracking their own statements, and there is too much cognitive demand required to track their previous statements and to sufficiently elaborate a statement they are currently making (Colwell et al., 2002, 2007). They are working to avoid disclosing sensitive information, making contradictions, or any other behaviors that could lead to a loss of credibility (e.g., lack of eye contact, admission of possible mistakes). The result is often less overall detail and a dearth of additional detail from deceivers throughout the interview.

The aforementioned sometimes explains the existence of systematic differences between honest and deceptive responding regarding a witnessed or experienced event. However, all of these differences are predicated upon appropriate interviewing, and no system of credibility assessment will ever exist without careful consideration to interviewing. A good interviewer must obtain unbiased information from honest respondents while exploiting the differences between honest and deceptive responding to facilitate the detection of deception. In other words, interviews are needed that facilitate honest recall while hindering and highlighting attempts to control information and to impression manage.

Investigative Interviewing

The goals of an investigative interview are (1) to maximize the amount of information obtained, (2) to minimize contamination of memory, (3) to generate statements that can be used in credibility assessment and (4) to maintain the integrity of the investigative process (Yuille, Hunter, Joffe, & Zaparniuk, 1993). These goals are listed in order of importance. This means that the primary consideration during an

investigation is to interview in a manner that obtains maximal accurate information from an honest respondent. Only after a framework has been developed that does this can an investigator or researcher attempt to implement strategies to discriminate honesty from deception. Detecting deception is pointless unless one first has created a mechanism to support honesty. To reiterate, an investigative interview is primarily a mechanism to gather information. Judging the veracity of that information is only meaningful to the extent that the information has been obtained in a manner that protected the memory of honest respondents.

There are a number of investigative interviews in existence. The Step-Wise Interview (SI; Zaparniuk et al., 1995; Colwell et al., 2002), for example, was created by Yuille for systematic assessment of statements and protection of memory. There is some research evidence to indicate that the SI does not work as well as the Cognitive Interview (CI; Fisher et al., 1989) and the Reality Interview (RI; Colwell et al., 2002) in the detection of deception, but is a very good assessment tool for obtaining statements where accuracy of information is paramount. Further research is necessary to determine the relative ability of the SI, CI, and RI in detecting deception. The SI is an excellent interviewing strategy when accuracy of information is paramount and detection of deception is not the goal (Colwell et al., 2002). No interview is appropriate for all situations. The SI is best for those situations where accuracy of information obtained is the most important consideration. The CI is best for those situations where maximizing the amount of information obtained is the most important consideration. Finally, the RI is best for those situations in which detecting deception is the most important consideration. The SI and the CI can be used to detect deception, but this is not their primary reason for existing. Similarly, the RI elicits accurate information from honest respondents, but the reason this interview exists is to facilitate the detection of deception. It is up to the interviewer to choose the most appropriate interview for each situation (Colwell et al., 2002).

The CI (Fisher et al., 1989; Memon, Meissner, et al., 2010; Memon, Zaragoza, Clifford & Kidd, 2010) is the oldest of the three and provided the basis for the RI (Colwell et al., 2002).[2] An early and important step of any investigative interview is the development of rapport (Walsh & Bull, present volume; Yarbrough et al., present volume; Yuille et al., 1993). Without rapport, it will not be possible to obtain complete and accurate information from an honest witness, thereby making the investigation less fruitful overall and hindering any attempts to detect deception (Colwell et al., 2002, 2007, 2008; Vallano & Compo, 2011).

The actual investigation portion of the interviewing begins with the elicitation of a free narrative. This allows for an honest person to provide information with as little potential for contamination as possible, and it has some valuable consequences for deceivers which are explained later. Following the free narrative, an interviewer can use mnemonics to enhance the respondent's recall for the event (Table 11.3).

[2] The Reality Interview was called the Inferential Interview in its original article (Colwell et al., 2002). However, many readers thought that *inferential* meant *untrained*, in that the group *inferred* their own style of interviewing. This was not correct, and the name was changed to avoid later confusion.

Table 11.3 Script for Reality Interview as used with students suspected of stealing an exam key

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

[a]The segment of the interview will later be used to guide scoring. Information from the open-ended questions is divided into that information obtained during Free Recall and information obtained during the Mnemonics. This allows for isolation of the recall enhancement effects of the Mnemonics

This approach is a direct form of the CI, and the field is grateful to the seminal research done by Fisher, Geiselman, and, later, Memon on this topic. Following the use of mnemonics, the respondent is asked to provide his or her description one last time. This last task, "Tell me everything again, and provide everything you remember even if you think it is irrelevant," can be considered a mnemonic in its own right. The basic structure of the RI is shown in Table 11.3. This table shows the script for the RI and how dependent variables are broken down as being elicited either by the free recall or by the mnemonics.

Interpersonal Dynamics and Recall Enhancement

As indicated above, an important early step of an investigative interview is the development of rapport. Sometimes, rapport requires the demonstration of empathy. It is important to mention that it is ethically questionable to demonstrate empathy in many forensic settings (Melton, Petrila, Poytherss, & Slobogin, 2007). Empathy is a powerful tool that should be used only when one is acting in a manner that is beneficial to the person being assessed. This means that empathy is acceptable during an investigative interview, where the goal is to gain information to ascertain what happened. Empathy is not acceptable during an interrogation, where the goal is to get a person to confess to a crime (Buckley, 2006). This is standard training for those in forensic psychology (Melton et al., 2007), but it is not something that appears to be widely known among investigators or psychologists who train investigators (Inbau, Reid, Buckley, & Jayne, 2004). In fact, participation in interrogation is not ethical for psychologists. Psychologists are to "do no harm" and are not to engage in activities that diminish the overall perception of psychology as a field. Psychologists who study investigative interviewing must walk a fine line and would do well to remember the drastic difference between interrogation and investigation. The safest approach is to develop rapport without the use of empathy. After all, empathy is necessary in therapy, but it is not necessary for an investigation.

The strength of rapport between the interviewer and the respondent is a primary factor in determining the amount and quality of information obtained during an investigative interview (Vallano & Compo, 2011). Most respondents need to be comfortable and to feel safe with the interviewer. Importantly, there must be a *transfer of control* to the respondent. This means that the respondents are taught that, in some respects, they are to lead the interview and to proceed at their own pace. They are instructed to take as much time as necessary to prepare a response, and they are informed that the interview process is meant to facilitate their responding. It is not meant to be a question-and-answer session controlled by the interviewer. To do this properly, it is good practice to have the respondent describe a couple of neutral events prior to discussion of the target event. In this way, the interviewer can teach the respondent about the process of the interview and what his or her responsibilities are. These descriptions also provide an opportunity to increase rapport. They have been considered as baselines for verbal behavior, but this is questionable as the sample of behavior obtained from a neutral event might be different from a sample of behavior obtained regarding the target event in an investigation (Colwell et al., 2007).[3] If done properly, the respondent will feel as comfortable as possible and will be aware that it is his or her responsibility to lead the interview (Colwell et al., 2002; Memon, Meissner, et al., 2010). This provides honest respondents with an environment that maximizes the utility of the mnemon-

[3] This chapter is concerned with verbal behavior, and this statement regarding difficulties in obtaining baselines during recall of innocuous events only applies to verbal behavior, and not to nonverbal behavior.

ics, and it creates for deceptive respondents an environment that highlights their attempts at control of information and impression management (Colwell et al., 2008; Suckle-Nelson et al., 2010).

Mnemonics and Recall Enhancement

A mnemonic is a memory aid. There are two general types of mnemonics: those that assist with encoding and those that assist with recall. Most witnesses and victims do not have advanced warning that they are about to experience something that will need to be remembered. Therefore, the mnemonics that aid in encoding are of minimal use to investigative interviewers. However, mnemonics that assist with recall are a tremendous asset to the investigative interviewer. The mnemonics used to enhance recall are based upon the principles of encoding specificity and spreading activation. Encoding specificity is the principle that any stimulus that was encoded at the time of a target event can serve as a retrieval cue for the memory of the target event (Fisher et al., 1989). Spreading activation is the notion that recalling parts of a target memory enhance one's ability to recall the remainder of that memory. Activation of one area of a memory network can facilitate the activation of other areas of that network. Practically, this means that the act of recall can become a positive-feedback system (Colwell et al., 2007).

The mnemonics used in the interview techniques discussed in this chapter are taken directly from early CI research. The first mnemonic to be used with the CI is *mental reinstatement of context* (Colwell et al., 2002; Fisher et al., 1989). This is an image-based technique in which the interviewer asks the respondent to think back to the time of the original event. The respondent is instructed to think of details from each sensory modality, as well as to describe his or her thoughts and feelings. He or she is also told to report everything even if he or she does not think it is important (another mnemonic). This is critical to the outcome of an interview. The mental reinstatement of context mnemonic can protect a respondent against subsequent contamination of memory (e.g., the "Geiselman effect"; Verkampt & Ginet, 2010). This could serve to partially inoculate against later misinformation and protects the memory trace (Memon, Zaragoza, et al., 2010). The second specific mnemonic is *recall from another perspective*. This attempts to get beyond the filtering effects of a respondent's schema for the target event. The respondent is asked, for example, to imagine if someone else had been in the room or to describe the event as someone else would have seen it. The third specific mnemonic is *reverse-order recall*. This is quite difficult for respondents but is very useful, especially for the detection of deception (Colwell et al., 2007, 2008, 2012; Vrij et al., 2006). Respondents are literally asked to describe the entire event but to begin with the end and end with the beginning. Finally, respondents are asked to retell the entire event, one last time.

In the context of an investigative interview, a mnemonic is a memory enhancement strategy used at the time of recall (Fisher et al., 1989). There is an interaction between mnemonics and honesty vs. deception that is vital to interviewing that

facilitates the detection of deception. Honest respondents are free to think about and completely report the target memory. Deceptive respondents are not, and instead must focus on their lie script. Because of this, honest respondents benefit from mnemonics to a higher degree than deceptive respondents (Colwell et al., 2002, 2007, 2008). Mnemonics, in general, lead to recall enhancement, but there is a difference in the recall enhancement for honest respondents compared to deceptive respondents. The variables that help honest people remember actually make the act of reporting and impression management more difficult for deceptive respondents. Mnemonics help honest respondents and hinder deceptive respondents; stated another way, mnemonics lead to DRE (Colwell et al., 2012).

Differential Recall Enhancement

The central lesson of the authors' last 16 years of research is: mnemonics and forced-choice questions enhance the reporting of honest respondents, allowing them to provide longer, more detailed, and spontaneously structured statements. These same mnemonics and forced-choice questions make responding more difficult for deceptive respondents, causing them to provide shorter, less detailed, and less spontaneously structured statements. There are two reasons for this DRE. First, a properly administered interview helps honest respondents to remember and to provide statements with a significant amount of additional words and details. Second, the same interview causes deceptive respondents to work harder and to rely more on their short, carefully phrased lie scripts (Ansarra et al., 2011; Colwell et al., 2002, 2007; Suckle-Nelson et al., 2010). Concretely, this DRE is manifested as the presentation of new information as a result of the mnemonics. Accordingly, the information provided during an investigative interview can essentially be divided into two phases – information provided prior to the use of mnemonics and additional information provided as a result of the mnemonics. In the CI and RI, this division is described as Free Recall (i.e., information presented before the mnemonics) and Mnemonics (i.e., additional information provided as a result of the mnemonics). Therefore, DRE can be highlighted by assessing the information provided during Free Recall vs. the information provided during the Mnemonics (Colwell et al., 2008, 2012).

DRE depends on a proper interview structure, appropriate mnemonics, and the operationalization of criteria suggestive of honesty or deception in a manner that takes advantage of this structure and content. Specifically, to take advantage of DRE, one must (1) obtain an original free narrative, (2) proceed with mnemonics and associated tasks, and finally, (3) perform a content analysis of the information derived with dependent variables divided across the free recall and mnemonics sections of the interview (see "Interview Portion for Scoring" column in Table 11.3). An effective exercise to demonstrate the DRE effect is to analyze the data obtained from an investigative interview in two ways. First, consider all of the information provided as a whole; that is, simply examine the content criteria of interest and calculate an average

Table 11.4 Accuracy of decisions assessing whole statement versus assessing free recall and mnemonics separately: Improvement from considering the DRE effect of the interview

	Accuracy of decisions (%)		
Citation	Whole statement	Free recall vs. mnemonics	Improvement to highlight DRE
Colwell et al. (2002)			
Transcribed verbal accounts from CI and RI	68.6	92.4	23.8
Colwell et al. (2007)			
Experiment 1			
Hand-written statements from RI	67.5	81.0	13.5
Colwell et al. (2007)			
Experiment 2			
Transcribed verbal accounts from RI	67.5	95.0	27.5
Colwell et al. (2008)			
Transcribed verbal accounts from RI	63.7	86.8	23.1
Suckle-Nelson et al. (2010)			
Transcribed verbal accounts from RI, males	76.5	88.3	11.8
Transcribed verbal accounts from RI, females	79.5	89.8	10.3
Average			18.3

value for the entire description. Second, take the same description, but examine the information obtained as a function of when it was first provided. Calculate a value for each content criterion based upon the free recall portion of the interview, and calculate a second value for each content criterion based upon the mnemonics portion of the interview. This gives the ability to assess the information provided at free recall, and then assess any new information that was provided as a result of the mnemonics. This second approach highlights the differential effects of the mnemonics for honest and deceptive respondents (i.e., highlights DRE). As seen in Table 11.4, based on our previous research, the second approach yielded an increase in the ability of the statement content criteria to discriminate between honest and deceptive statements.

Comparing the Cognitive Interview to Reality Interview

The CI and the RI are two interviews that are formulated to have both the structure and the content necessary for DRE. The two-alternative, forced-choice questions of the RI are included to facilitate the detection of deception. The first block of these

forced-choice questions begins after the participant has completed the mental reinstatement of context mnemonic. This placement is crucial. The mental reinstatement of context mnemonic helps inoculate the participant's memory against contamination (Memon, Meissner, et al., 2010; Memon, Zaragoza, et al., 2010), and the order of the RI is designed to take advantage of this protective effect (Colwell et al., 2002, 2008). Forced-choice questions should *not* be used prior to the elicitation of two free narratives: the first from a general, open-ended free recall and the second from the mental reinstatement of context mnemonic. They should require the respondent to think deeply about the event (e.g., visuospatial questions are best), and they should be about factors that the respondent is not likely to have practiced as part of his or her lie script. Inferences (e.g., Was the gun closer to the door or to the window? Did anyone intend to harm anyone else?) require more cognitive effort than simple recall (Colwell et al., 2002), thereby maximizing differences between honest and deceptive respondents.

A major challenge with forced-choice questions is to avoid leading the respondent. There are two general strategies to do this. In those cases where a significant amount of information is not available at the time of the interview (which is likely the case in most preliminary investigative interviews), questions should require very general inferences. Examples of these are, "Was there a crime committed," or, "Did anyone speak with an accent." When information is available regarding the target event prior to the interview, the forced-choice questions can be global inferences, or they can be carefully constructed so that one of the choices is absolutely correct.

The basic task of certain types of deception is to inhibit memory for the original event while providing information from a lie script in a manner that avoids contradictions and appears confident. The forced-choice inferences interfere with this inhibition of the original memory in two ways: (1) they provide information from the target event as one of the response choices whenever possible and (2) they force respondents at least to think outside their script and, at best, to think back to the target event. The first is akin to a Stroop Task, where the automatic tendency to process information from the target event will compete with the effortful attempt to suppress that information. The second is simply another form of an unanticipated question.

There will be times when the respondent is deceptive but has no memory for the target event. In these cases, forced-choice questions increase the cognitive demand placed upon respondents because they must choose carefully while attempting to determine what the interviewer does and does not know about the event. In instances in which a significant amount of information about the target event is known, the forced-choice questions can provide an additional cue to credibility, in a manner akin to symptom validity testing. Simply put, people who are being deceptive often perform at or below chance, or at least significantly worse than what should be expected, indicating that they are deliberately missing questions to manipulate the interviewer (e.g., Colwell & Colwell, 2011; Colwell & Sjerven, 2005; Hiscock & Hiscock, 1989; Rogers & Bender, 2003). Recent research also indicates that forced-choice questions can be used to screen a large number of witnesses to focus on those

who either are very cooperative or are working hard to hide information (Gavigan et al., 2012).

The authors have compared the RI to the CI in two studies using male and female inmates who witnessed a staged theft. All participants witnessed the theft. Those in the honest group were asked to describe what they had seen and cooperate in the investigation of the thief. Those in the deceptive group were instructed to answer questions in such a way that investigators would not be able to convict the real thief. In the first study, the dependent measures were TTR, response length, and coherence. Results indicated a ceiling effect, with both interviews performing in the mid-90% range in accurately classifying statements as honest or deceptive (93% for RI, and 94% for the CI; Colwell et al., 2002).

In the second study, the open-ended narratives elicited by free recall and each of the mnemonics were assessed, and the dependent measures were the amount, type, and location of details (Hiscock-Anisman et al., 2012). The actual answers to the forced-choice questions were not considered in this study. Therefore, raters coded what appeared to be identical interview formats (note: the forced-choice questions were omitted from the transcripts). The RI led to significant improvement over the CI in predictive accuracy (90% vs. 71%). This means that 90% of RI statements were accurately classified as honest or deceptive, while only 71% of the CI statements were accurately classified as honest or deceptive. The amount of information provided at free recall was the same for both interviews, which was expected. However, during the mnemonic phase of the interview, honest respondents in the RI provided more detail than did honest respondents in the CI. Also, deceptive respondents in the RI provided less detail than did deceptive respondents in the CI. This study demonstrated that RI is better able to generate DRE. The forced-choice questions made deception more difficult and obvious while providing yet another memory cue for honest respondents. The differences between the CI and RI were primarily in the form of the amount of words and details added during the mnemonic section of the interviews. There was not a corresponding difference in how carefully phrased statements became during the mnemonic section as measured by the TTR. This, along with a ceiling effect, appears to be why there was no difference in predictive accuracy of the CI vs. the RI in the original Colwell et al. (2002) study.

Assessment Criteria Indicative of Deception: Combining Differential Recall Enhancement with Content Analysis

The Assessment Criteria Indicative of Deception (ACID; Colwell et al., 2008) system integrates interviewing for DRE with dependent measures that highlight vividness and spontaneity for honest respondents, and highlight careful phrasing and control of information for deceptive respondents (Colwell et al., 2007, 2012). Optimally, the ACID approach uses an RI to elicit the statement; a CI, structured as above, can be used, but this has been shown to be less effective (Hiscock-Anisman et al., 2012). The dependent measures for the ACID system are response length,

TTR, amount of details, coherence,[4] and whether the respondent admitted that he/she could possibly be mistaken. Response length, TTR, and the number of details are tallied for free recall, and then separately for the mnemonics phase of the interview. Response length is simply the total number of words provided in the statement. TTR is scored by computer software, first for the free recall phase and then averaged across responses to the mnemonics phase. Finally, details are tallied as the number of specific descriptors used in the statement. In order for a detail to be counted in the mnemonics phase, it must be unique—that is, it must not have been provided during the free recall phase. Only new details are counted.

As seen in Table 11.5, this system has been used successfully with university students, male and female prison inmates, children, people speaking English and Arabic, and US military personnel who experienced severe anxiety and distress as part of their training. ACID has been used to study the statements of victims, witnesses, and perpetrators. Finally, ACID can also be applied to written transcripts of interviews, instant messenger interactions over the Internet, and audio statements assessed in real time.

Perhaps the best example of the utility of ACID was a study examining statements provided by college students regarding the theft of an exam key (Colwell et al., 2008). University students were required to enter what they believed to be a professor's office and steal or replace what they thought was an exam key. Students were told that the professor who used the office did not know of the study and, if they were caught, the professor would be angry. It was also stated that the police would be called, and the student would be arrested, and would have to wait until either the Department Chair or the Principal Investigator could come and explain things for them (note: this was a deception, but students reported that they believed this part of the experiment during debriefing). After completing the illicit act, participants were assigned to either report honestly (i.e., answer completely and help the investigator) or deceptively (i.e., answer so that they are not found guilty of anything). Participants were also offered $100 for the "two most convincing" statements. The students had approximately one week to practice their statements prior to returning for their interview. The interview followed the RI format provided in Table 11.3. The only answer from the forced-choice questions that was analyzed was whether the participants admitted that they could have been mistaken. The other dependent measures were the number of details provided during free recall (i.e., external-free recall, contextual-free recall, and internal-free recall), the number of new details added during the mnemonics (i.e., external-mnemonics, contextual-mnemonics, and internal-mnemonics), and the total number of words provided during the mnemonic section of the interview. On the basis of these eight variables, 86.8% of statements were accurately classified as honest or deceptive (78.9% of honest and 94.7% of deceptive statements were accurately classified). Honest statements were longer and more detailed during the mnemonics, and more likely to

[4] The authors suggest using an expanded version of coherence with children. For children, whether they disclose sensitive information should be scored as a "yes or no." In addition, the number of serious contradictions should be counted.

Table 11.5 Studies using all or part of the ACID system

Citation	Statement type	How decisions were made	Accuracy of decisions
Colwell et al. (2002)	Inmate transcribed statements, staged theft w/ live actors	TTR, response length verbal hedges, coherence	91.9%
	Interviewed with CI and RI	Discriminant function analysis	
Colwell et al. (2007), Experiment 1	Student hand-written statements, videotaped theft, videotaped interviewer	Amount and distribution of details	81.0%
	Interviewed with CI		
		Discriminant function analysis	
Colwell et al. (2007), Experiment 2	Inmate transcribed statements, staged theft w/ live actors	Amount and distribution of details	95.0%
	Interviewed with RI		
		Discriminant function analysis	
Colwell et al. (2008)	Student transcribed statements, theft from professor's office	Amount and distribution of details, admitting mistakes	86.8%
	Interviewed with RI		
		Discriminant function analysis	
Colwell et al. (2009)	Transcripts from students and inmates from previous studies	Student rater judgments	
		Pre-ACID training	57.0%
		Post-ACID training	77.0%
Suckle-Nelson et al. (2010)	Inmate transcribed statements, staged theft w/ live actors	TTR, response length, amount and distribution of details	
	Interviewed with RI	Males	88.3%
		Females	89.8%
		Discriminant function analysis	
Morgan, Hazelett, Colwell (2011)	Military personnel statements in high-stress advanced training	TTR, response length, amount and distribution of details	82.0%
	Interviewed with CI		
		Discriminant function analysis	

(continued)

Table 11.5 (continued)

Citation	Statement type	How decisions were made	Accuracy of decisions
Ansarra et al. (2011)	Student statements re: theft from classroom Combined sample, CI and RI	Amount and distribution of details	80.0%
		Discriminant function analysis	
Montalvo et al. (in press)	Written or audio statements from previous research	Student rater judgments after 3-h ACID training	
		Untrained audio	51.0%
		Untrained written	55.0%
		Trained audio	71.0%
		Trained written	71.0%
Hiscock-Anisman, Colwell et al. (2012)	Written or audio statements from previous research	Police officer rater judgments after 8-h ACID training	
		Untrained	56.0%
		Trained	90.0%
Colwell, Hiscock-Anisman et al. (2011)	Arabic-speaking community residents who stole from hotel room Interviewed with CI	TTR, response length, amount and distribution of details	83.3%
		Discriminant function analysis	
Colwell, Colwell et al. (2011)	Written transcripts from previous studies	Forensic professional rater judgments after 3-h ACID training	
		Pre-training	58.2%
		Post-training	70.0%
Williams et al. (2012)	Children's honest or deceptive statement re: scavenger hunt Interviewed with RI	Amount and distribution of details, number and location of contradictions / sensitive disclosures, admitting mistakes	87.0%
		Discriminant function analysis	
Hiscock-Anisman, Morrissey et al. (2012)	Inmate-transcribed statements, staged theft w/ live actors	Response length, amount and distribution of details	
		RI	
		CI	90.0%
		Discriminant function analysis	71.0%

Note: In all of the above-mentioned training studies, participants were trained to locate and track differences between honest and deceptive statements previously obtained. This was a demonstration of how people can be trained to perform credibility assessment and detecting deception within the ACID framework. Learning to interview, which is the most important part of ACID, requires a 2–4-day training session.

contain an admission of a potential mistake, compared to deceptive statements. Consistent with DRE, the most powerful predictors were obtained during the mnemonic segment of the interview. This study demonstrated the utility of ACID with a student sample engaged in what was believed to be an illicit act. Deceivers were able to either omit certain information or to otherwise modify an existing memory during the investigative interview. This is arguably the most common type of deception that is encountered in real investigations and has, therefore, been the most common type of situation studied in ACID research (Colwell et al., 2002, 2008; Colwell, Hiscock-Anisman, et al., 2006; Suckle-Nelson et al., 2010). It is, after all, considerably simpler to tell a partial truth than it is to wholly fabricate.

Another important scenario facing investigators is a deceptive respondent who is wholly or largely fabricating; that is, describing an event they have not actually witnessed or performed. In one such study, US military personnel were asked to either respond honestly or deceptively about undergoing torture and interrogation as part of their training (Morgan, Hazlett, & Colwell, 2011). Honest respondents had undergone torture and interrogation as part of their training, whereas deceptive respondents were military personnel who were qualified for this same training, but who had not been through the process. Rather, deceptive respondents were provided with a description taken from the Internet that had been posted by someone who went through the training, and were asked to respond as if they had been through the same. In this study, a CI was used to elicit statements, which were analyzed on the basis of the TTRs, response length, and the amount of detail at free recall and the amount of new detail added during the mnemonics. This allowed for an 82% rate of accurately classifying statements as honest or deceptive. Honest statements had lower TTRs, longer responses, and more detail. Again, the largest effects were seen during the mnemonic section of the interview.

Moderators of Assessment Criteria Indicative of Deception

Gender

The highest predictive accuracies in ACID research to date have occurred when the sample is either all male or all female (or split by gender) and when statistical software makes the predictions using a discriminant function analysis. These differences due to gender were verified in a recent study by Suckle-Nelson et al. (2010). This study demonstrated that women who responded honestly were able to provide more information than were men who responded honestly. Also, women who responded deceptively were more aware of the need to keep their statement short and careful than were men who responded deceptively. Importantly, men who responded honestly were more likely to provide an incoherent story than were women who responded deceptively. Research has shown that women in the US tend to have improved attention, memory, interpersonal, and verbal ability compared to men (Crawford, 1995). Those differences are likely the partial cause of these observations.

Language

Initial research using ACID has shown that the technique demonstrates significant success with English speakers. All and parts of the system have also been used to discriminate honest from deceptive statements from Arabic speakers. In these studies, a sample of Arabic speakers was questioned through an interpreter, and the interpreter's English translation of the Arabic speaker's responses was coded (Colwell, Hiscock-Anisman, Hazlett, & Morgan, 2011). As seen in Table 11.5, ACID was able to discriminate honest from deceptive responding at an 83.3% level of accuracy. It is important to note that this study employed a CI rather than an RI because the CI is more user-friendly. Further research should employ the RI to generate a more accurate estimate of the ability of ACID through an interpreter.

The authors attempted to use ACID with Chinese speakers, using pictograms rather than verbal statements to score the dependent variables. This study was promising in that the strategies of impression management described by Chinese respondents did not differ in many ways from the strategies of impression management listed in Table 11.2 (Hsieh et al., 2012). However, the stimulus chosen for the event was faulty, and the use of pictograms may have also been inappropriate. The stimulus for this research was faulty because Chinese and US students were each asked to either respond honestly about a time someone with authority mistreated them, or to lie and make a false-allegation that someone with authority had mistreated them. All the US students in the honest group had experiences where a professor, teacher, parent, or coach had mistreated them, and they were willing to disclose. Similarly, the US students did not have any difficulties making false allegations. However, the authors learned from communication with the Chinese scholars that Chinese people are generally taught that any incident where it appears an authority figure is mistreating someone represents a misperception, and a chance for personal growth on behalf of the person who thought he or she was mistreated. In this study, the authors asked the participants to respond honestly about a situation that, in their culture, does not exist. Additionally, ACID variables were created to be scored on written words. However, because the strategies of impression management and deception that were described by the Chinese participants were the same as those described by the US participants, it is likely that ACID will work with the Chinese statements. Future research should give careful attention to the event chosen (e.g., one that does not violate cultural assumptions of the participants) and also score the ACID content criteria using audio rather than written statements.

It is important to note that ACID, or part of ACID, has been used in the assessment of people whose first language was Spanish but who were speaking English (Colwell et al., 2002, 2007; Suckle-Nelson et al., 2010), that is, because these samples were drawn from Texas prisons and approximately 10% of participants spoke Spanish before learning to speak English (Colwell, 1997). Yet, there was no difference in the ability to detect deception in any of these studies as a function of ethnicity. Also, ACID research has included English speakers from across the US, Canada, and Scotland, Arabic speakers from Morocco, and Chinese speakers from

China. The technique worked well with all of these samples except for the flawed Chinese study. Even in that study, the strategies described by participants regarding their attempts at impression management and deception were essentially the same as the strategies described by a matched sample of US college students. This convergence of verbal behavior and strategies of deception indicates that ACID may be assessing basic aspects of interpersonal deception regarding an event. After all, DRE is based upon memory and cognition, and these should be common regardless of ethnicity or culture. It may be possible to synthesize a uniform theory of interpersonal deception using these and related findings.

Training and Modality

Perhaps the most impressive aspect of ACID research is the recent findings related to training and application. These include the following: (1) ACID can be trained to a significant degree with a half-day workshop (e.g., participants are able to improve from chance to the 70% range), (2) ACID can be applied to real-time audio recordings rather than just written transcripts and (3) a one-day training workshop is sufficient to improve police officers' ability to detect deception from either transcribed or audio statements from chance levels to almost 90% success (Hiscock-Anisman et al., 2012). To date, almost all statement analysis systems have required many days of training and have been done using verbatim transcripts. These are tedious and make application difficult, at best. However, a series of studies has shown that ACID can be easily trained, with no difference in the ability of those trained to detect deception by reading or by listening to statements. This ability ranges from the mid-70% range following a half-day training to 90% following a full-day training (Colwell & Colwell, 2011; Hiscock-Anisman et al., 2012; Montalvo et al., in press). In fact, ACID can even be applied to the statements obtained via instant messenger in computer-mediated interactions (Werdin et al., 2012).

All of these training studies have one very important feature – the statements provided for making judgments of honest vs. deceptive all come from unique events. The participants (e.g., college students, forensic professionals, police officers) were each presented with a number of honest or deceptive statements. Each statement was the only one given about a particular event. This means that each participant would have two statements from witnesses regarding thefts, two statements from suspects regarding what they did during the time of two different alleged thefts, and two statements from respondents who allege that they were mistreated by their boss or professor. Each statement had to be judged on its own merit as there was no other evidence regarding the event described in each. Importantly, participants were not able to compare the descriptions provided on one statement with the description of the same event provided on another. The findings from training studies have indicated that a full day is better than a half day (Colwell et al., 2009, 2012; Hiscock-Anisman et al., 2012) and that decisions can be made just as well from audio or written statements (Kradas et al., 2012; Montalvo et al., in press).

Applications, Future Research, and Limitations

One area of emerging interest is computer-mediated communication (see Hancock & Woodworth, present volume). This area has been of interest to the authors since the beginning of their research. Colwell's (1997) research, for example, used software to score TTR and response length. This led to a natural grouping among the ACID dependent measures – TTR and response length are scored by computer (and admitting potential mistakes could easily be scored by computer), whereas the amount of detail presented at free recall and the amount of new detail presented during the mnemonics is scored by trained raters. There is a natural tendency in automated applications to emphasize TTR and response length, and there is a natural tendency in interpersonal interactions to just use the amount and type of details. Future research should be done to compare the validity of each of these simplified approaches. There is some reason to believe that the two may work as well as one another (Morgan et al., 2011), although some loss is likely in predictive accuracy from using less content criteria (Suckle-Nelson et al., 2010). This loss of accuracy may be outweighed by the ability to listen to a real-time audio of an interview and make a decision regarding honesty vs. deception in the 88–90% range (Hiscock-Anisman et al., 2012).

One of the most challenging areas for investigative interviewing and statement analysis is communication via instant messenger. This challenge also allows for potential insight into the component processes of interpersonal deception. ACID has been based upon the findings that deceptive respondents work harder than honest respondents due to the need to (1) track information and avoid releasing sensitive details or making contradictions and (2) appear calm and confident in the interpersonal setting. Instant messenger interactions provide the chance to review and edit prior to sending, and allow for one to see the history of the interaction. Similarly, there is no face-to-face interaction, so there is less behavior to control (see Hancock & Woodworth, present volume). Werdin et al. (2012) studied instant messenger interactions obtained from men or women who were trying either to tell the truth about their gender or lie about their gender. Honest respondents described the last time they did something with their same-sex best friend, while deceptive respondents were required to fabricate an interaction as if they were the other gender and were spending time with their best friend of that same gender. These participants went through the standard ACID technique. It was possible to accurately classify 30 of 37 statements as honest or deceptive. Importantly, honest statements were longer and more detailed at free recall, but they were not significantly longer or more detailed during the mnemonics. The differences were all in the expected direction, but they were not significant. Nevertheless, these findings could give some insight into the process of deception. The authors are currently trying to replicate this study, and are also studying what will happen if mirrors and cameras are placed in front of the respondents. If the expected differences return as a result of re-introducing the video information, it will underscore the amount of effort expended by deceivers in trying to appear calm and confident. Similarly, another

variation will be investigated where respondents are unable to see their response history and so will have to track information with no additional cues. This manipulation will provide insight into the amount of effort expended in working memory during deception.

The principle of DRE, in general, and the ACID system, in particular, can apply to any situation where an honest respondent should have formed an episodic memory. This means that it can be applied to eliciting and assessing information about what happened during a certain period of time. An example of an area that could see future applications is the assessment of Post Traumatic Stress Disorder (e.g., regarding allegations of abuse). In contrast, this approach will not work when questioning people about their attitudes. It is also not likely to work when questioning about future plans (unless you can have a person describe what he or she has done to prepare for his or her future behavior). Similarly, these techniques are not likely to assist in detecting deception about what a person may be hiding in his or her clothing. Finally, this approach is not likely to work when the respondent actually believes what he or she is saying. The latter is an interesting empirical issue, and future researchers would do well to consider whether this type of assessment ceases to assist in detecting deception when the deceiver comes to believe his or her deception is true. Related to this, DRE and ACID are not likely to work when people are mistaken. This system is not designed to detect memory errors, only deliberate deception.

The most important area to study at this point is real-world application. There have been a number of lab-based studies, but no, "real world," evidence to date. This type of research is expensive and difficult because it requires cooperation with an investigative agency. Moreover, it requires that investigators are willing to seriously apply these techniques in their own work, rather than the techniques they have used for the entirety of their career. Professionals in law enforcement and forensics have been trained in ACID. The response has been positive. However, there is still no available data regarding systematic application by professional investigators.

Summary and Conclusions

There has been a large body of research over the past two decades dealing with the importance of proper interviewing in order to obtain information and detect deception. This has led to a focus on those techniques that help honest respondents remember and provide information, while hindering attempts at deception. This interaction effect between question type and honesty of responding is DRE. Many researchers are working on variations of this, including those studying cognitive load, strategic or tactical interviewing, and ACID. ACID is a systematic approach to interviewing and assessment with content criteria derived from CBCA, RM, and research into interpersonal deception and impression management. ACID derives statements using either the CI or the RI, with the RI being preferred when detecting deception is the primary goal of the interview. ACID has been studied in numerous

settings, and it can be trained to students and professionals alike. It has been used with English interpretations of Arabic statements. Also, English speakers from the US, Canada, and Europe, as well as Chinese speakers from China and Arabic speakers from Morocco all approach the process of interpersonal deception in the same way. Therefore, it appears that DRE and ACID are getting at basic aspects of human interactions that apply across cultures. This technique is most capable when dealing with face-to-face interviews and statements involving episodic memory. Other types of deceptions are outside the ability of this technique. Some areas of future research are computer-mediated communication and PTSD. This line of research has potential to inform actual investigative interviews, as well as provide insight into the process of interpersonal deception.

Acknowledgement The authors thank Lori Colwell, Ph.D., for her comments on previous versions of this chapter.

References

Akehurst, L., Koehnken, G., & Hofer, E. (1995). Content credibility of accounts derived from live and video presentations. *Legal and Criminological Psychology, 6*, 65–83.

Ansarra, R., Colwell, K., Hiscock-Anisman, C., Hines, A., Fleck, R., Kondor, S., & Cole, L. (2011). Augmenting ACID with affective details to assess credibility. *European Journal of Psychology Applied to the Legal Context, 3*(2), 1–10.

Blandon-Gitlin, I., Pezdek, K., Lindsay, S., & Hagen, L. (2009). Criteria-based content analysis of true and suggested accounts of events. *Applied Cognitive Psychology, 23*(7), 901–917.

Blandon-Gitlin, I., Pezdek, K., Rogers, M., & Brodie, L. (2005). Detecting deception in children: An experimental study of the effect of event familiarity on CBCA ratings. *Law and Human Behavior, 29*(2), 187–197.

Boychuk, T. D. (1991). *Criteria-based content analysis of children's statements about sexual abuse: A field-based validation study*. Unpublished doctoral dissertation. Arizona State University: Tempe, Arizona.

Buckley, J. P. (2006). The Reid Technique of interviewing and interrogation. In T. Williamson (Ed.), *Investigative interviewing: Rights, research, regulation* (pp. 190–206). Devon, UK: Willan Publishing.

Colwell, K. (1997). *Interviewing techniques and the psycholinguistic assessment of statement credibility*. Unpublished master's thesis. Sam Houston State University: Huntsville, Texas.

Colwell, L. H., & Colwell, K. (2011). Assessing feigned cognitive impairment in defendants hospitalized for competency restoration: Further validation of the TOMI. *Journal of Forensic Psychology Practice, 11*, 293–310.

Colwell, L. H., Colwell, K., Hiscock-Anisman, C., Hartwig, M., Cole, L., Werdin, K., & Youschak, K. (2012). Teaching professionals to detect deception: The efficacy of a brief training workshop. *Journal of Forensic Psychology Practice, 12*, 68–80.

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

Colwell, K., Hiscock-Anisman, C. K., Corbett, L., Bonilla, Y., Memon, A., & Hauselt, W. J. (2011). Change in suspect's memory as a result of deception. *American Journal of Forensic Psychology, 29*(4), 1–9.

Colwell, K., Hiscock-Anisman, C.K, Hazlett, G., & Morgan, C.A. (2011). *Credibility assessment of suspects speaking Arabic through English interpreters*. Paper presented at the Annual Conference of the American Psychology and Law Society, Miami, Florida.

Colwell, K., Hiscock-Anisman, C., Memon, A., Colwell, L. H., Taylor, L., & Woods, D. (2009). Training in Assessment Criteria Indicative of Deception (ACID) to improve credibility assessment. *Journal of Forensic Psychology Practice, 9*, 199–207.

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. K., Memon, A., Taylor, L., & Prewett, J. (2008). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*, 167–180.

Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. (2006). Strategies of impression management among deceivers and truth-tellers: How liars attempt to convince. *American Journal of Forensic Psychology, 24*(2), 31–38.

Colwell, L. H., Miller, H. A., Miller, R. S., & Lyons, P. M. (2006). US police officers' knowledge regarding behaviours indicative of deception: Implications for eradicating erroneous beliefs through training. *Psychology, Crime, & Law, 12*(5), 489–503.

Colwell, K., & Sjerven, E. R. (2005). The 'Coin-in-Hand' strategy for forensic assessment of malingering. *American Journal of Forensic Psychology, 23*, 83–89.

Crawford, M. (1995). *Talking difference: On gender and language*. Thousand Oaks, CA: Sage Publications, Inc.

Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police to interview tactically. *Journal of Investigative Psychology and Offender Profiling, 8*(2), 189–202.

Esplin, P.W., Boychuk, T.D., & Raskin, D.C. (1988). *A field validity study of criteria-based content analysis of children's statements in sexual abuse cases*. Paper presented at the NATO Advanced Study Institute on Credibility Assessment, Maratea, Italy.

Fisher, R., Geiselman, E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of the actual victims and witnesses of crime. *Journal of Applied Psychology, 74*(5), 722–727.

Gavigan, B.J., Fede, J., Richards, E., Hiscock-Anisman, C.K., Pankratz, L., & Colwell, K. (2012). *A forced-choice test to detect witness deception*. Poster presented at the Annual Conference of the American Psychology and Law Society, San Juan, Puerto Rico.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(4), 643–659.

Hartwig, M., & Doering, N. (2009). *Strategies of the deception game: A social cognitive theory of the psychology of innocence and guilt*. Paper presented at the annual conference of the American Psychology-Law Society, San Antonio, Texas.

Hartwig, M., Granhag, P., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*(5), 603–619.

Hernández-Fernaud, E., & Alonso-Quecuty, M. (1997). The cognitive interview and lie detection: A new magnifying glass for Sherlock Holmes? *Applied Cognitive Psychology, 11*(1), 55–68.

Hines, A., Colwell, K., Hiscock-Anisman, C., Garrett, E., Ansarra, R., & Montalvo, L. (2010). Impression management strategies of deceivers and honest reporters in an investigative interview. *European Journal of Psychology Applied to the Legal Context, 2*(1), 73–90.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11*(6), 967–974.

Hiscock-Anisman, C. K., Colwell, K., Danna, M., Rodriguez, N., Sorcinelli, A., & French, M. N. (2012). *Using ACID to help police detect deception*. Poster presented at the Annual Conference of the American Psychology and Law Society, San Juan, Puerto Rico.

Hiscock-Anisman, C.K., Morrissey, D., Willet, S., Evans, M., Belarde, D., & Colwell, K. (2012). *Comparing the Reality Interview and the Cognitive Interview: Two-alternative, forced-choice questions and differential recall enhancement*. Manuscript in preparation.

Hofer, E., Akehurst, L., & Metzger, G. (1996). *Reality monitoring: A chance for further development of CBCA*? Paper presented at the annual meeting of the European Association of Psychology and Law, Sienna, Italy.

Hsieh, C., Hiscock-Anisman, C.K., Colwell, K., Florence, S., Sorcinelli, A. & French, M.N. (2012). *Strategies of impression management and deception among Chinese and English speakers*. Paper presented at the Annual Conference of the American Psychology and Law Society. San Juan, Puerto Rico.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2004). *Criminal interrogations and confessions* (4th ed.). Sudbury, MA: Jones & Bartlett Publishers, Inc.

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

Karim, A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., & Birbaumer, N. (2010). The truth about lying: Inhibition of the anterior prefrontal cortex improves deceptive behavior. *Cerebral Cortex, 20*(1), 205–213.

Koehnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80*(6), 671–684.

Kradas, M., Henry, C., Williams, S., Talwar, V., Marsland, K.W., & Colwell, K. (2012). *Reliability and validity of Assessment Criteria Indicative of Deception (ACID) with audio and written statements*. Manuscript in preparation.

Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Hershkowitz, I. (1997). Criteria-based content analysis: A field validation study. *Child Abuse & Neglect, 21*, 255–264.

Landry, K., & Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability to detect deception in adults. *Law and Human Behavior, 16*, 663–675.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585–589.

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, & Law, 11*(1), 99–122.

Melton, G. B., Petrila, J., Poythress, N. G., & Slobin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: The Guilford Press.

Memon, A., Fraser, J., Colwell, K., Odino, G., & Mastroberadino, S. (2009). Distinguishing truthful from invented accounts using Reality Monitoring criteria. *Legal and Criminological Psychology, 15*(2), 177–194.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, & Law, 16*(4), 340–372.

Memon, A., Omerod, T.C., & Dando, C.J. (2012). *Truth or lies: Reality monitoring and deception detection*. Manuscript under review.

Memon, A., Zaragoza, M., Clifford, B. R., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law and Human Behavior, 34*(2), 105–117.

Montalvo, L., Hallinan, C., Hiscock-Anisman, C.K., Morrissey, D., Bonilla, Y., Colwell, K., & Kradas, M. (in press). Using ACID to improve credibility assessments with written and audio statements. *American Journal of Forensic Psychology.*

Morgan, C. A., Hazlett, G., & Colwell, K. (2011). Efficacy of forensic statement analysis in distinguishing truthful from deceptive eyewitness accounts of highly stressful events. *Forensic Sciences, 56*(5), 1227–1234.

Porter, S., & Yuille, J. C. (1995). Credibility assessment of criminal suspects through statement analysis. *Psychology, Crime, & Law, 1*(4), 319–331.

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*(4), 443–458.

Rogers, R., & Bender, S. D. (2003). Evaluation of malingering and deception. In A. M. Goldstein (Ed.), *Handbook of psychology: Forensic psychology* (Vol. 11, pp. 109–129). Hoboken, NJ: John Wiley and Sons, Inc.

Ruby, C. L., & Brigham, J. C. (1998). Can criteria-based content analysis distinguish between true and false statements of African American speakers? *Law and Human Behavior, 22*, 369–388.

Scheck, B., Neufeld, P., & Dwyer, J. (2001). *Actual innocence: When justice goes wrong and how to make it right*. New York, NY: Signet.

Steller, M., Wellershaus, P., & Wolf, T. (1988). *Empirical validation of criteria-based content analysis*. Paper presented at the NATO Advanced Study Institute on Credibility Assessment, Maratea, Italy.

Suckle-Nelson, J. A., Colwell, K., Hiscock-Anisman, C., Florence, S., Youschak, K. E., & Duarte, A. (2010). Assessment Criteria Indicative of Deception (ACID): Replication and gender differences. *Open Criminology, 3*, 23–30.

Undeutsch, U. (1954). *Die Entwicklung der gerichtspsychologischen Gutachtertätigkeit*. Oxford, England: Hogrefe, Verlag Fuer Psychologie.

Vallano, J. P., & Compo, N. S. (2011). A comfortable witness is a good witness: Rapport building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology, 25*(6), 960–970.

Verkampt, F., & Ginet, M. (2010). Variations of the cognitive interview: Which one is the most effective in enhancing children's testimonies? *Applied Cognitive Psychology, 24*, 1279–1296.

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, & Law, 11*(1), 3–41.

Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior, 26*(3), 261–283.

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.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences, 10*(4), 141–142.

Vrij, A., Leal, S., Granhag, P., Mann, S., Fisher, R.P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior*, *33*(2), 159–166.

Werdin, K., Colwell, K., Hiscock-Anisman, C.K., Hartwig, M., Bessenoff, G., & Fede, J. (2012). *ACID and computer-mediated deception: The use of Assessment Criteria Indicative of Deception to assess credibility via instant messaging*. Manuscript under review.

Williams, S.M., Talwar, V., Rodriguez, N., Fede, J., Colwell, K., & Hiscock-Anisman, C.K. (2012). *The mistakes of child lie tellers: Using the assessment criteria indicative of deception (ACID) system with children*. Poster presented at the Annual Conference of the American Psychology and Law Society, San Juan, Puerto Rico.

Winkel, F. W., & Vrij, A. (1995). Verklaringen van kinderen in interviews: Een experimental onderzoek naar de diagnostische waarde van Criteria Based Content Analysis. *Tijdschrift voor Ontwikkelingspsychologie, 22*, 61–74.

Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 95–115). New York, NY: Guilford Press.

Zaparniuk, J., Yuille, J., & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry, 18*(3), 343–352.