A comparison of different investigative interviewing techniques in generating differential recall enhancement and detecting deception

Document status and date:
Published: 01/01/2019

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Download date: 25 Apr. 2024

RESEARCH ARTICLE

WILEY

A comparison of different investigative interviewing techniques in generating differential recall enhancement and detecting deception

Jacob De Rosa[1] | Cheryl Hiscock-Anisman[2] | Adam Blythe[1] |
Glynis Bogaard[3] | Ashley Hally[1] | Kevin Colwell[1] (iD)

[1] Department of Psychology, Southern Connecticut State University, New Haven, Connecticut

[2] Professor of Psychology, The National University, La Jolla, California

[3] Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands

Correspondence
Kevin Colwell, PhD, Professor of Psychology, Department of Psychology, Southern Connecticut State University, 501 Crescent Street, New Haven, CT 06515.
Email: colwellk2@southernct.edu

Abstract

This study compared the Stepwise Interview, Cognitive Interview, and Reality Interview in detecting deception with inmates. The dependent measures were the amount of unique details provided during the free narrative and mnemonics and the number of words provided during the free narrative and mnemonics of each interview. The Stepwise Interview generated 58.3% accuracy, the Cognitive Interview generated 70.0% accuracy, and the Reality Interview generated 93.3% accuracy. The different tasks of these interviews increased the differences between honest and deceptive statements and therefore, increased the accuracy in detection of deception. Differential recall enhancement is used to explain the findings.

KEYWORDS

credibility assessment, detecting deception, investigative interviewing

1 | INTRODUCTION

Detecting deception is difficult for professionals and laypersons alike (Colwell, James-Kangal, Hiscock-Anisman, & Phelan, 2015; Hauch, Sporer, Michael, & Meissner, 2016; Mann, Vrij, & Bull, 2004; Porter & Brinke, 2010). However, certain interview techniques can enhance recall for honest respondents yet increase the difficulty experienced by deceptive respondents. The concept of differential recall enhancement (DRE) was created to explain the performance

There are no known conflicts of interest regarding this work.

- -

of these techniques (Colwell et al., 2015; Colwell, Hiscock-Anisman, & Fede, 2013). Thus, interviews that generate more DRE lead to larger differences between honest and deceptive statements and improve the detection of deception[1] (Ansarra et al., 2011; Colwell, Hiscock, & Memon, 2002; Colwell et al., 2009; Colwell, Hiscock-Anisman, Memon, Rachel, & Colwell, 2007; Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Hartwig, Granhag, Strömwall, & Vrij, 2005; Liu et al., 2010; Suckle-Nelson et al., 2010).

The research into DRE began with the recognition of the different cognitive and interpersonal demands placed upon deceptive versus honest respondents (Colwell et al., 2002; Colwell et al., 2007; Mann & Vrij, 2006; Vrij, Granhag, Mann, & Leal, 2011). For example, deceptive respondents are required to provide enough information to satisfy the interviewer without making contradictions or releasing sensitive information (Ansarra et al., 2011; Colwell et al., 2013; Colwell, Hiscock-Anisman, Memon, Taylor, & Prewett, 2008; Hartwig, Granhag, & Strömwall, 2007; Hines et al., 2010; Tekin et al., 2015; Tekin, Granhag, Stromwall, & Vrij, 2016).

Deceptive respondents often form a lie script to cope with these demands (and those who do not prepare a script in advance often attempt to use their first description of the target event as a script). A deceptive respondent can think about and respond from this script rather than memory for the original event. The script provides a way to satisfy the interviewer without the worry of releasing sensitive information or making contradictions. Thus, a lie script is a powerful tool for impression management (Colwell et al., 2007; Colwell et al., 2013; Porter & Yuille, 1996; Suckle-Nelson et al., 2010). Honest respondents do not perceive the same need for control of information and impression management. Therefore, honest respondents are free to respond based upon their memory of the original event without the preparation of a script (Colwell, Hiscock-Anisman, Memon, & Michlik, 2006; Hines et al., 2010). For honest respondents, there is less concern with minor contradictions or deviations from previous descriptions of the target event. Lie scripts mean that deceptive respondents are attempting to stick to the same story and tell it the same way each time. Honest respondents, who are less likely to use scripts, are also less likely to worry about changes.

Interviewers can take advantage of this situation. An open-ended recall task that allows a deceptive respondent to present a lie script and an honest respondent to present a free narrative is the first step (Colwell et al., 2002). The next steps employ mnemonics and careful recall tasks that interfere with a deceptive respondent's ability to inhibit her or his memory for the original event. These mnemonics take advantage of the different cognitive demands of honest versus deceptive responding (Colwell et al., 2007; Colwell et al., 2013), increase the cognitive load placed upon deceptive respondents (Vrij, Leal, Mann, & Fisher, 2012), and highlight attempts at impression management on behalf of deceptive respondents (Colwell et al., 2013; Hartwig, Granhag, Stromwall, & Doering, 2010). The mnemonics that have been used previously are mental reinstatement of context, recall from another perspective, and reverse-order recall. The careful recall tasks used previously have been of the two-alternative, forced-choice format. These present a choice between a correct and an incorrect piece of evidence from the target event. If there is not enough evidence that is known with certainty, then these questions may require a, "yes or no," response about some aspect of the original event that is not likely to be part of a prepared script (Colwell et al., 2013).

In clinical assessment, two-alternative, forced-choice questions are frequently used on their own, such as in the assessment of malingering (Colwell & Colwell, 2011; Colwell, Colwell, & Perry, 2008; Colwell & Sjerven, 2005; Hiscock & Hiscock, 1989). In an investigative interview, the presentation of two-alternative, forced-choice questions begins after several precautions are taken. These include rapport, transfer of control to the respondent, the elicitation of a free narrative, and a mental reinstatement of context mnemonic so that the respondent's memory is protected as much as possible from contamination (Colwell et al., 2002; Colwell et al., 2013; Memon, Zaragoza, Clifford, & Kidd, 2010).

There are three widely used investigative interviews that contain some, or all, of the steps listed above. They differ in process in a manner that should elicit different amounts of DRE. The Stepwise Interview (Table 1; Porter, Juodis, & Leanne, 2010; Yuille, 1989) comprises rapport-building, transfer of control, and the elicitation of a free-

---

[1]Credibility assessment determines whether a statement appears to be an honest recall of a genuine memory. Detection of deception determines whether a statement is honest versus deliberately distorted. In instances such as the present work, when the ground truth is known, credibility assessment becomes truth confirmation and detecting deception remains recognising lies. The techniques studied in this paper are tools for both credibility assessment and detection of deception.

TABLE 1 Stepwise Interview

Interview segment	Recall task	Scoring segment
Demographics and rapport	Discussion with RA. Simple conversation.	None
Recall of two events	1. Please describe, in as much detail as possible, everything you can remember about your first day on this Prison Unit. 2. Please describe, in as much detail as possible, everything you can remember about the last meal you have eaten.	None
Free narrative	"" everything you can remember about the event that you witnessed.	Free narrative
Probe 1—Research	"" everything you can remember about the research assistant.	Mnemonics
Probe 2—Participant	"" everything you can remember about the man with the newspaper.	Mnemonics
Probe 3—Thief	"" everything you can remember about the last man to enter the room.	Mnemonics
Retelling	This is the last question. Will you, "" everything you can remember about the event that you witnessed?	Mnemonics

Note. "" = "Please describe, in as much detail as possible."

narrative prior to more specific recall tasks (Colwell, Hiscock-Anisman, & Fede, 2014; Vallano & Compo, 2013). The Cognitive Interview (Table 2; Geiselman, 1984; Memon, Meissner, & Fraser, 2010) begins with the exact same tasks as the Stepwise Interview. The difference happens when the Cognitive Interview employs open-ended mnemonics designed to elicit additional free narrative responses, where the Stepwise Interview moves into specific, open-ended, probes. The mnemonics utilise multiple retrieval pathways, spread of activation, and schema-consistent recall to enhance recall for honest respondents. These same mnemonics take advantage of the differing cognitive and inter-personal demands of deception by using non-directive, open-ended tasks, increased cognitive demands/cognitive load for deceivers (Colwell et al., 2002; Colwell et al., 2007; Vrij, Fisher, Mann, & Leal, 2008). The Reality Interview (Table 3; Colwell et al., 2013; Colwell et al., 2002; Colwell et al., 2015) is identical to the Cognitive Interview, but with the addition of the forced-choice questions. These present additional recall cues for honest respondents. They also make it more difficult for a deceptive respondent to inhibit memory of the original event and/or to use a lie script.

TABLE 2 Cognitive Interview

Interview segment	Recall task	Scoring segment
Demographics and rapport	Discussion with RA. Simple conversation.	None
Recall of two events	1. Please describe, in as much detail as possible, everything you can remember about your first day on this Prison Unit. 2. Please describe, in as much detail as possible, everything you can remember about the last meal you have eaten.	None
Free narrative	"" everything you can remember about the event that you witnessed.	Free narrative
Mental reinstatement of context	Think back to the event you just described. Take as long as you would like. Think about the sights, sounds, smells, the other people in the room, and your own thoughts and feelings. Take your time. When you are ready, please describe everything you can remember.	Mnemonics
Recall from another perspective	Imagine that you had been somewhere else in the room (i.e., in another seat, or in the place of another person). "" everything you would have seen from that perspective.	Mnemonics
Reverse order recall	"" everything that you can remember in reverse order, beginning with the last thing you witnessed and ending with the first.	Mnemonics
Retelling	This is the last question. Will you, "" everything you can remember about the event that you witnessed?	Mnemonics

Note. "" = "Please describe, in as much detail as possible."

TABLE 3 Reality Interview

Interview segment	Recall task	Scoring segment
Demographics and rapport	Discussion with RA. Simple conversation.	None
Recall of two events	1. Please describe, in as much detail as possible, everything you can remember about your first day on this Prison Unit. 2. Please describe, in as much detail as possible, everything you can remember about the last meal you have eaten.	None
Free narrative	"" everything you can remember about the event that you witnessed.	Free narrative
Mental reinstatement of context	Think back to the event you just described. Take as long as you would like. Think about the sights, sounds, smells, the other people in the room, and your own thoughts and feelings. Take your time. When you are ready, please describe everything you can remember.	Mnemonics
Forced Choice 1	1. Did you witness a crime? 2. Was this an act of violence? 3. If a police officer had been there, would anyone have been arrested?	Not scored
Recall from another perspective	Imagine that you had been somewhere else in the room (i.e., in another seat, or in the place of another person). "" everything you would have seen from that perspective.	Mnemonics
Forced Choice 2	4. Did anyone speak with an accent? 5. Did anyone intend to harm anyone else? 6. Have you ever witnessed anything like this before?	Not scored
Reverse order recall	"" everything that you can remember in reverse order, beginning with the last thing you witnessed and ending with the first.	Mnemonics
Forced Choice 3	7. Did anything about the event you witnessed seem unusual? 8. Which was closer to the front of the desk, the clock or the rings? 9. Do you think you could be mistaken about anything you have said so far?	Not scored
Retelling	This is the last question. Will you, "" everything you can remember about the event that you witnessed?	Mnemonics

Note. "" = "Please describe, in as much detail as possible."

Finally, forced-choice questions model brief and non-detailed responding, which may entice deceptive respondents into providing short responses (Ansarra et al., 2011; Colwell, Hiscock-Anisman, et al., 2008; Colwell et al., 2002). These interviews should create a continuum in their ability to elicit DRE. The smallest differences between honest and deceptive statements should be created by the Stepwise Interview. Larger differences between honest and deceptive statements should be created by the Cognitive Interview. The largest differences between honest and deceptive statements should be created by the Reality Interview. Similarly, the Stepwise Interview should correspond with the lowest accuracy rate in detecting deception. The Cognitive Interview should generate a significantly higher accuracy rate. The Reality Interview should generate the highest rate of accuracy in detecting deception.

Current research

No investigative interview is optimal for all situations. The interview used in each specific situation should be chosen based upon the respondent and the purpose of the interview (Colwell et al., 2002). The current effort compared the

Stepwise Interview, the Cognitive Interview, and the Reality Interview in their ability to elicit DRE and facilitate the detection of deception. A mock theft using live actors was staged in front of incarcerated male witnesses. Each participant was interviewed in one of the three conditions. The hypotheses were based upon DRE. The dependent measures were chosen because of their ability to demonstrate recall enhancement—the amount of information and the number of words both are excellent ways to operationalise recall performance. The entire ACID system was not used because the other variables in the system focus upon impression management (TTR) and upon contradictions within a statement or between a statement and available evidence (coherence). These variables were not as relevant to DRE.

Increased DRE should correspond to more words and unique details added after a free narrative for honest respondents, and fewer words and unique details added after a free narrative by deceptive respondents. Therefore, it was hypothesised that (a) the Stepwise Interview would not elicit significant differences in the number of words or number of details added when comparing honest with deceptive statements. (b) The Cognitive Interview would elicit significant differences in the number of words and number of details added after the free narrative when comparing honest with deceptive statements. (c) The Reality Interview would elicit significantly more words and unique details added in the statements of honest respondents and would elicit significantly fewer words and details added in the statements of deceptive respondents when compared with the Cognitive Interview. (d) The Stepwise Interview was hypothesised to have the lowest classification accuracy. (e) The Cognitive Interview was hypothesised to have an intermediate level of classification accuracy. (f) The Reality Interview is hypothesised to have the highest level of classification accuracy.

2 | METHOD

2.1 | Participants

This study examined statements provided by 180 incarcerated male participants. About 56.7% of participants endorsed "Caucasian," as their ethnicity, 25.6% endorsed "African American," 14.4% endorsed "Hispanic," 1.6% endorsed "Asian," and 1.7% endorsed "other." These participants were selected by sending a written description of the study to every fifth name on the general prison roster, excluding those who were in administrative segregation.

Incarcerated participants were chosen because this guaranteed that each participant had experienced an investigative interview. Also, the research was designed to be relevant to police interviewing, and this sample matches those who are interviewed by police much better than a college student sample. Finally, inmates have demonstrated a better awareness of deception and its detection than other groups; thus, it was important to make certain that techniques derived studying college students worked the same with inmates (Hartwig, Granhag, Strömwall, & Andersson, 2004). The prison was a general population unit, which meant that the participants were convicted of an array of offences, including drug crimes, assault, robbery, and sex offences. It was not possible to collect data regarding the actual crime of record for each participant.

A total of 228 participants were presented the scene and interviewed. Of these, 180 were randomly selected (60 per interview type) for transcription, coding, and inclusion in this data set.

2.2 | Design

The participants were randomly assigned in a 3 × 2 experimental matrix. The independent variables were interview (stepwise, cognitive, and reality) and honesty (honest responding vs. deceptive responding). There were 30 participants in each of the six cells, so that $N = 180$.

2.3 | Procedure

2.3.1 | Staged theft

After arrival and informed consent, the participants were seated in a desk in a classroom. During this time, there was a research assistant (RA), a Caucasian woman in her early 30s, dressed in a lab coat, seated behind the desk at the front of the classroom. When everyone was quiet and facing the front of the room, the scene began. The RA took her rings off, placed them on the desk, and applied lotion to her hands. At this point, the imposter participant entered the room. He was a Caucasian man in his early 20s of notable stature. He was 6'7" tall and weighed about 285 lb. The imposter participant was obviously not one of the inmates due to the way he was dressed and his briefcase. However, he went through the same basic sequence of behaviours as the participants. He entered the room and asked if this was the "memory study." He was told he was in the correct place. Then he was given a consent form that he was told to have a seat, read, and sign the form. The imposter participant took a seat beside and somewhat in front of the RA's desk, facing the real participants. Once seated, he opened his briefcase, removed a newspaper, and began reading it.

The thief barged in and began yelling about waiting and being locked out. He was a Caucasian man in his late 20s, just over 6 ft tall, who weighed about 200 lb. The RA told the thief to take a seat and then read and complete the informed consent document. The RA looked at the time, said she had to speak with the professor, and rushed out of the room. The thief stood behind the desk. He surveyed the items on the desk. The he looked at the crowd and smiled. He deliberately and obviously stole several of the items. The real participants had a clear view of the desk and the thief, but the imposter participant was focused on his newspaper. The thief stole several items and put them into his coat pockets. Among these, he stole four rings and put them on his fingers (they did not fit properly). Finally, he stole a piece of gum, unwrapped it, and put it into his mouth. He smiled at the audience again. Then the thief told the imposter participant that he had another appointment, and he left. The RA returned, looked at her desk, and realised that things were gone. This same scene was repeated as closely as possible for small groups of participants, ranging from 13 to 27. There were no major deviations from the scene with any of these groups (the researchers/actors had performed this same scene more than 20 times each by the time of this experiment).

2.3.2 | Instructions and motivation

At this point, the participants were given randomised packets that determined their interview and honesty conditions. Each participant met with a RA who reviewed his instructions with him. Each participant was told that he was to participate in an investigative interview regarding the theft that he had witnessed. Those in the honest condition were instructed to "Report everything as honestly and completely as possible, making certain that the Thief can be found guilty." Those in the deceptive condition were instructed to "Imagine that you know the Thief. Lie so that the Thief will not be found guilty." It is not ethically permissible to provide monetary incentives to incarcerated participants. Therefore, motivation was addressed through (a) attempting to create a strong interviewer–interviewee relationship within a comfortable setting and (b) the pep talk. In the pep talk, a member of the research team spoke with each witness for about 2 min. All were reminded that the research aimed to move law enforcement towards information gathering and away from manipulation or interrogation. Those in the honest condition were reminded of the importance of being able to remember and describe what they had seen (i.e., told that the research would never have an impact on the way that police treat victims, witnesses, and suspects unless the respondents worked hard to provide a complete and accurate memory). Those in the deceptive condition were told of the importance of providing a convincing lie and making certain that the thief was not held accountable (i.e., told that the research would never have an impact on how police treat victims, witnesses, and suspects unless the deceivers worked hard to deceive the interviewer and escape detection).

2.3.3 | Interviews

Participants were given approximately 5 min to prepare for the interview. Each participant was interviewed by one of nine trained undergraduate interviewers using either a Stepwise Interview, Cognitive Interview, or Reality Interview. All interviewers were trained in each interview.

Interviewers were trained by the professor who runs the lab. The training involved a 1 hr lecture and viewing of a video of each interview. The interviewers then interviewed one another. Finally, each interviewer recorded a video of themselves using each interview script. The professor viewed each video with the interviewer from that video and provided feedback on technique. At this point, the interviewer was cleared to do interviews for the study.

These statements were transcribed and prepared for analysis. The entire interaction was transcribed. These transcripts were first examined to determine that interviews had been conducted properly. The forced-choice segments of the Reality Interview were redacted so that each transcript appeared to have the same number of segments. Following this, a scoring version was created that only showed the statements of the participant. The raters did not know which interview elicited each statement.

2.3.4 | Scoring

The dependent measures were Number of Unique Details Free Narrative (DFN), Number of Unique Details added during Mnemonics (DMN), Response Length Free Narrative (RFN), and the Response Length Mnemonics (RMN). Each detail was only counted the first time it was presented. Those initially presented during the Free Narrative were tallied as DFN. Those initially presented during the interview segments that followed the Free Narrative were tallied as DMN. DFN and DMN were scored by hand. RFN and RMN were scored by Wordscan software, which has an established test–retest reliability of 1.0 (Colwell et al., 2002). Details were counted only if they were relevant. Also, each detail was only counted once. As an example, the sentence "He placed the wallet on the desk" has a new detail for almost every word. "He," "placed," "wallet," "on the," and "desk" are each unique details. This gives a total of 5. Now, assume that in the same statement, the sentence, "He placed the red wallet on the wooden desk," was presented. This sentence would only contain two new details, "red" and "wooden."

Undergraduate raters were provided standard training. Raters were given a lecture about how to recognise relevant statement details. Then they were provided with four transcripts from previous studies. The first transcript was scored as a group. The second transcript was scored with two raters working together. The remaining transcripts were scored by each rater working individually. Each rater's scoring of the fourth transcript was compared with a key. Those raters who could not agree with the key regarding DFR and DMN at above a 90% level were required to score yet another transcript. This training resulted in 10 raters, with an average proportion agreement to the coding key of 93%. One additional rater scored RFR and RMN using Wordscan.

Each rater scored 36 transcripts, and each transcript was scored twice, by two different raters. Rater 1 scored Transcripts 1–18, Rater 2 scored Transcripts 18–1, and this reverse-order pairing was continued until Rater 9 scored Transcripts 161–180 and Rater 10 scored transcripts 180–161. Each pair of raters was given six transcripts from the Stepwise Interview, six transcripts from the Cognitive Interview, and six transcripts from the Reality Interview. ICCs were calculated to demonstrate interrater reliability. ICC values between raters were 0.86 for DFN and 0.89 for DMN.

3 | RESULTS

The first manipulation check was the review of instructions with each participant. Second, each transcript was read by two raters and separated into two groups based upon whether the statement contained in each could be used as evidence against the alleged thief. This grouping was compared with the experimental instructions. The two raters

agreed on all statements,[2] thereby demonstrating that all participants followed their instructions. Third, it was verified that each statement was elicited using the proper interview. Finally, each participant was asked to endorse whether they had responded honestly or deceptively as a part of a post-interview questionnaire. There was 100% agreement between these endorsements and the honesty condition.

Each of the dependent measures generated skewness and kurtosis statistics between −2 and +2. There was no need to assess homogeneity of variance due to having 30 participants per cell. Therefore, the data were considered acceptable for further analysis.

The first analysis for this study was an overall (2 × 3) MANOVA to determine whether there were significant differences in the dependent measures of DFN, DMN, RFN, and RMN as a function of interview and honesty. There was a significant effect for honesty, $V = 0.58$, $F(7,171) = 33.63$, $p < 0.001$, $\eta_p^2 = 0.58$ and a significant interview × honesty interaction, $V = 0.39$, $F(14,340) = 5.97$, $p = 0.008$, $\eta_p^2 = 0.20$. The statement characteristics (means and standard deviations) elicited by each interview, and the differences between honesty and deception within each interview are presented in Table 4.

3.1 | Interaction effect

The different amount of DRE elicited by each interview created an interview × honesty interaction in the number of details and words obtained. This interaction was examined for each interview, for both honest and deceptive respondents (Table 5).

3.1.1 | Honest respondents

For honest respondents, there was a significant relationship between interview and DMN; $F(2,87) = 18.73$, $p < 0.001$, partial eta^2 = 0.30. An LSD test indicated that the mean difference in DMN between the Stepwise Interview and the Cognitive Interview of 15.27 (95% CI = 7.20, 23.34) was significant, $p = 0.037$. The mean difference in DMN between the Cognitive Interview and the Reality Interview of 10.18 (95% CI = 1.96, 18.40) was also significant, $p = 0.025$. Thus, for honest respondents, the Stepwise Interview elicited the fewest number of DMN, the Cognitive Interview elicited an intermediate number of DMN, and the Reality Interview elicited the largest number of DMN.

For honest respondents, there was a similar effect of interview upon RMN; $F(2,86) = 11.5$, $p < 0.001$, partial eta^2 = 0.21. An LSD test indicated that the mean difference in RMN between the Stepwise Interview and the Cognitive Interview of 63.15 (95% CI = 21.89, 104.42) was significant, $p = 0.006$. The mean difference in RMN between the Cognitive Interview and the Reality Interview of 38.31 (95% CI = 3.71, 80.33) was also significant, $p = 0.024$. Thus, for honest respondents, the Stepwise Interview elicited the fewest number of RMN, the Cognitive Interview elicited an intermediate number of RMN, and the Reality Interview elicited the largest number of RMN.

3.1.2 | Deceptive respondents

For deceptive respondents, there was a significant relationship between interview and DMN; $F(2,86) = 12.70$, $p < 0.001$, partial eta^2 = 0.22. An LSD test indicated that the mean difference in DMN between the Stepwise Interview and the Cognitive Interview of 12.08 (95% CI = 2.60, 21.57) was significant, $p = 0.008$. The mean difference in DMN between the Cognitive Interview and the Reality Interview of 10.99 (95% CI = 1.65, 20.32) was also significant, $p = 0.003$. Thus, for deceptive respondents, the Stepwise Interview elicited the largest number of DMN, the Cognitive Interview elicited an intermediate number of DMN, and the Reality Interview elicited the lowest number of DMN.

---

[2]One participant misunderstood the instructions and chose to lie to get the thief out of trouble by talking about the effect that conditions of confinement have upon one's judgement. This case was dropped and replaced by one of the reserve cases that were assessed in anticipation of data attrition.

TABLE 4 Descriptive statistics and effect sizes for DV's comparing honest versus deceptive statements within each interview

	Stepwise Interview						
	Honest		**Deceptive**				
	Mean	**SD**	**Mean**	**SD**	**F (1,59)**	**p**	η_p^2
DFN	47.92	17.98	42.77	10.56	1.95	0.167	0.03
DMN	47.40	12.46	47.90	12.46	0.02	0.903	<0.01
RFN	307.20	39.74	125.43	40.56	5.16	0.031	0.08
RMN	310.03	61.95	293.4	62.70	1.07	0.303	0.02
	Cognitive Interview						
	Honest		**Deceptive**				
	Mean	**SD**	**Mean**	**SD**	**F (1,59)**	**p**	η_p^2
DFN	45.91	18.25	36.38	10.82	5.66	0.024	0.09
DMN	62.51	17.11	35.55	26.68	22.46	0.003	0.28
RFN	294.41	131.50	112.93	50.85	7.31	0.008	0.11
RMN	373.19	105.24	225.34	112.07	28.23	0.005	0.32
	Reality Interview						
	Honest		**Deceptive**				
	Mean	**SD**	**Mean**	**SD**	**F (1,59)**	**p**	η_p^2
DFN	45.82	11.53	42.31	4.56	1.64	0.213	0.03
DMN	72.68	17.87	24.56	14.11	135.43	0.002	0.70
RFN	306.82	138.60	114.06	52.39	9.00	0.003	0.13
RMN	411.50	68.63	192.75	73.87	139.85	0.008	0.71

Note. DFN: detail free narrative; DMN: detail mnemonics; RFN: response length free narrative; RMN: response length mnemonics.

For deceptive respondents, there was a similar effect of interview upon RMN; $F(2,86) = 11.2$, $p < 0.001$, partial eta^2 = 0.20. An LSD test indicated that the mean difference in RMN between the Stepwise Interview and the Cognitive Interview of 68.6 (95% CI = 24.10, 112.01) was significant, $p < 0.001$. The mean difference in RMN between the Cognitive Interview and the Reality Interview of 32.59 (95% CI = 3.71, 80.33) approached significance with a one-tailed $p = 0.065$. Thus, for deceptive respondents, the Stepwise Interview elicited the largest number of RMN, the Cognitive Interview elicited an intermediate amount, and the Reality Interview elicited the lowest number of RMN.

3.1.3 | Summary of interaction effects

The Reality Interview elicited the most details (DMN) and words (RMN) of any interview from honest respondents and the fewest details and words from deceptive respondents. The Stepwise Interview elicited very similar numbers of details and words for both honest and deceptive respondents. The Cognitive Interview performed at an intermediate level, falling between the other two. The Reality Interview was the most successful at magnifying the differences between honest and deceptive respondents, that is, the Reality Interview elicited the most DRE.

3.1.4 | Accuracy rates

One of the most important aspects of the present study was the comparison of the interviews in classifying statements as honest or deceptive. A discriminant function analysis was used within each interview to make these classifications. These rates are presented in Table 6.

TABLE 5 Pairwise comparisons of the interview × honesty interaction effects.

Stepwise Interview versus Cognitive Interview				
Interaction effect	**Dependent measure**	**F (1,113)**	**p Value**	**η_p^2**
Interview × honesty	DFN	1.17	0.282	0.01
	DMN	17.89	<0.01	0.13
	RFN	0.01	0.921	<0.01
	RMN	16.56	<0.001	0.12
Cognitive Interview versus Reality Interview				
Interaction effect	**Dependent measure**	**F (1,113)**	**p Value**	**η_p^2**
Interview × honesty	DFN	1.01	0.317	0.01
	DMN	9.02	0.003	0.07
	RFN	0.10	0.752	<0.01
	RMN	4.47	0.037	0.04
Reality Interview versus Stepwise Interview				
Interaction effect	**Dependent measure**	**F (1,113)**	**p Value**	**η_p^2**
Interview × honesty	DFN	1.4	0.239	<0.01
	DMN	33.0	<0.001	0.22
	RFN	0.92	0.340	<0.01
	RMN	21.1	<0.001	0.15

Note. DFN: detail free narrative; DMN: detail mnemonics; RFN: response length free narrative; RMN: response length mnemonics.

TABLE 6 Classification rates as honest versus deceptive using DFA with DFN, DMN, RFN, and RMN as predictors

Stepwise Interview			
	Predicted honest (%)	**Predicted deceptive (%)**	**Overall percentage accuracy**
Honest	17 (56.7)	13 (43.3)	
Deceptive	12 (40.0)	18 (60.0)	
			58.3%
Cognitive Interview			
	Predicted honest (%)	**Predicted deceptive (%)**	**Overall percentage accuracy**
Honest	22 (20)	8 (10)	
Deceptive	9 (10)	21 (20)	
			70.0%
Reality Interview			
	Predicted honest (%)	**Predicted deceptive (%)**	**Overall percentage accuracy**
Honest	27 (90.0)	3 (10.0)	
Deceptive	1 (0.3)	29 (96.7)	
			93.3%

The binomial statistic was used to determine whether the interview accuracy rates differed from chance, as well as from one another. First, the accuracy rate generated by the Stepwise Interview (58%) was compared with the expected chance rate of 50%. For 60 observations, the number correct generated by the Stepwise Interview is 35. This was a statistically significant improvement over the number correct of 30 associated with chance performance, p = 0.045. Next, the accuracy rate generated by the Cognitive Interview (73%) was compared with the accuracy rate

generated by the Stepwise Interview (58%). For 60 observations, the number correct for the Cognitive Interview (43) was a statistically significant improvement in accuracy over the number correct of 35 generated by the Stepwise Interview, $p = 0.010$. Finally, the accuracy rate generated by the Reality Interview (93.3%) was compared with that generated by the Cognitive Interview (73%). For 60 observations, the number correct generated by the Reality Interview (56) was a statistically significant improvement over the number correct generated by the Cognitive Interview (43), $p < 0.001$. The Reality Interview led to the highest classification accuracy, followed by the Cognitive Interview, followed by the Stepwise Interview.

4 | DISCUSSION

This study demonstrated the relative ability of three investigative interviewing techniques. The Stepwise Interview, Cognitive Interview, and the Reality Interview were compared in eliciting information, eliciting DRE, and detecting deception. The Reality Interview performed better than the others in all three of these areas. These findings are made more important by the following: (a) The interviews and assessment criteria have a theoretical basis (Colwell et al., 2013). (b) The interviews are existing techniques that are not difficult for professionals (Colwell et al., 2012; Colwell et al., 2015) or lay people (Colwell et al., 2009; Montalvo et al., 2013) to learn and apply. (c) The content criteria are also easy to train and apply (Colwell et al., 2009; Colwell et al., 2015). (d) There are no special tools required. (e) The techniques can be used for face-to-face interactions, telephone conversations, and Internet-mediated communication (Colwell et al., 2013).

4.1 | Theoretical implications

The construct of DRE provides a theoretical mechanism that explains the relative performance of investigative interviews. For an interview to generate DRE, it must increase the amount of unique information provided by honest respondents while decreasing the amount of unique information provided by deceptive respondents. Honest statements grow through spontaneous addition of detail through the course of the interview, whereas deceptive statements tend to not have additional of detail through the course of the interview.

The DRE paradigm emphasises helping honest respondents to provide information. It assumes that the goal of research in investigative interviewing is to increase the accuracy of legal decision-making. To accomplish this, it is important to create situations in which honest witnesses and innocent suspects feel comfortable and have every chance to tell their version of events (Vallano & Compo, 2013). This is at least as important as increasing the amount of effort required by deceptive respondents. Catching liars can be a good thing, but finding the truth is almost always a good thing (in an investigation).

The construct of DRE applies to techniques other than those traditionally used in the ACID system (Colwell et al., 2013). For example, Strategic Use of Evidence (SUE; Luke et al., 2012) is a careful interviewing technique whereby evidence is told to a respondent after she or he has already been through much of the interview and has therefore committed to a version of the events. This additional evidence causes deceptive respondents to become more guarded and should reduce the amount of new information she or he is able to provide (SUE also increases or highlights statement–evidence inconsistencies). For honest respondents, the evidence that is strategically revealed provides a recall cue. The act of explaining what could initially appear as contradictory evidence leads to the addition of a significant amount of unique detail. Tactical interviewing works in much the same way (Dando & Bull, 2011). DRE simply provides a theoretical framework to speak about any technique that facilitates recall for honest respondents while making deception more difficult and more obvious.

4.1.1 | Types of statements

In this study, participants were free to choose their strategy of deception, within the limits set by the need to convince the interviewer of the suspect's innocence. They were given some time to prepare their lies as research has shown that liars usually prepare themselves before an interview (Hartwig et al., 2007).

A common strategy used by the deceptive respondents was to falsely accuse either the RA, the imposter participant, or one of the other inmate participants. By using this strategy, the deceptive respondent could be honest about most of the staged theft and only fabricate, omit, or distort details about who stole the items.

Even deceptive respondents who deliberately added information to the event from imagination still provided fewer details in response to mnemonics than honest respondents. Thus, deception was more than simply leaving out a few key details. It involved a careful style of speaking and close adherence to previously provided information. The care that participants gave to avoid contradictions, or the disclosure of sensitive information, speaks well to the motivation of the participants.

It is important to recognise that a lack of optimal or maximal motivation on behalf of participants does not negate these findings. The current study forms one piece of a larger puzzle. It should be evaluated in combination with studies that allowed for more incentives and higher stress. Students stealing an exam key from a professor's office (Colwell, Hiscock-Anisman, et al., 2008) and students stealing a wallet from a classroom (Ansarra et al., 2011) all found the same patterns of speech and detail as those found in the current study. In both of these studies, participants had the potential to obtain extra credit and $100 for success at theft and convincing the interviewer of their innocence (Ansarra et al., 2011; Colwell et al., 2008). The same patterns of speech and detail emerge when comparing honest versus deceptive statements for highly stressful events described by military personnel (Morgan, Colwell, & Hazlett, 2011).

4.2 | Practical applications

To be practical, a system of investigative interviewing and detection of deception must (a) obtain as much information as possible from honest respondents, (b) maximise differences between honest and deceptive responding, and (c) these differences need to be easy to observe. These concerns drove the selection of the dependent measures. The entire structure of these interviews—transfer of control to the respondents, an emphasis upon rapport, the use of practice recall tasks, the elicitation of a free narrative, and (in two of the interviews) the use of a mental reinstatement of context mnemonic—are intended to maximise recall and protect memory. It is relatively simple to have a computer count the number of words in a response, and it is relatively simple to teach an observer to count the number of details that a respondent adds during the mnemonic portion of an interview (Ansarra et al., 2011; Colwell et al., 2012; Colwell et al., 2015; Colwell et al., 2009; Montalvo et al., 2013).

The five studies cited above demonstrate that training people to detect deception by assessing the number of unique details and words provided to the segments of a Reality Interview is a relatively simple task. In each of these, a training of between 4 and 8 hr was able to significantly improve the ability to classify statements as honest or deceptive. The current data help explain why that is so. There were almost more 50 new details added by honest respondents compared with deceptive respondents following the free narrative. Honest respondents tended to add more new unique details after the free narrative than they provided during the free narrative. Deceptive respondents tended to add fewer unique details after the free narrative than they provided during the free narrative. Thus, training a professional to discriminate honest from deceptive responding using this system should be a rather straightforward task. First, they would listen to (or read) the free narrative and count the details presented. Then they would listen to (or read) the remainder of the interview and count the number of new details added. If the number of details added is more than the number initially provided, the statement is credible

and likely honest. If the number of details added is less than the number initially provided, the statement is not credible and likely deceptive.

An important note is that in the Reality Interview, the two dependent measures that were best at discriminating honesty from deception, DMN and RMN, have equal effect sizes (η_p^2 = 0.70 and η_p^2 = 0.71, respectively). It is suggested by the authors that these two criteria are used together when possible. However, there will be situations where one or the other presents itself more readily. For example, an electronic statement can easily be fed into Wordscan software. A recorded statement must be transcribed at present in order to use Wordscan, but this same recording can be assessed directly by those trained in detail analysis (Colwell et al., 2015; Montalvo et al., 2013).

4.2.1 | Limitations

The primary limitations of this study were the lack of major incentives for honest responding and the lack of negative consequences for deceptive respondents who were caught. The benefit of freedom versus the pain of incarceration creates a powerful social situation for those in actual investigations, and it is just not possible to match that intensity in a lab setting. In this study, for control reasons, each participant witnessed the same scene. In application, investigators are often speaking with only one or two witnesses/suspects regarding each event. It would be good for the field for researchers to also use designs in which each participant is describing a different event. However, this situation would create a difficulty in establishing the ground truth for each event.

In sum, the current study showed that a combination of mnemonics and forced-choice questions can trigger DRE and provide a very powerful tool for investigative interviewing and detecting deception with adults (Colwell et al., 2013; Colwell et al., 2015; Morgan et al., 2011; Morgan, Rabinowitz, Leidy, & Coric, 2014). The Reality Interview maximised DRE, as it was intended. The construct of DRE provides an interesting theoretical construct to describe and evaluate the performance of techniques designed to facilitate the detection of deception. Both the techniques and the underpinning theory should be used by future professionals in application and researchers in pursuit of understanding.

ORCID

Kevin Colwell [iD] https://orcid.org/0000-0003-1005-1699

REFERENCES

Ansarra, R., Colwell, K., Hiscock-Anisman, C., Hines, A., Fleck, R., Cole, L., & Belarde, D. (2011). Augmenting acid with affective details to assess credibility. *European Journal of Psychology Applied to the Legal Context*, *3*(2), 141–158.

Colwell, K., Colwell, L., & Perry, A. (2008). The test of malingered incompetence (TOMI): A forced-choice instrument for assessing cognitive malingering in competence to stand trial evaluations. American journal of. Retrieved from https://www.researchgate.net/profile/Kevin_Colwell/publication/234115494_tomi_ajfp_2008_electronic_article/links/09e4150f52aebd5e1c000000.pdf

Colwell, K., Hiscock, C., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology*, *16*(3), 287–300. https://doi.org/10.1002/acp.788

Colwell, K., Hiscock-Anisman, C., & Fede, J. (2014). Assessment criteria indicative of deception: An example of the new paradigm of differential recall enhancement. *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment.*. https://doi.org/10.1007/978-1-4614-5547-9_11

Colwell, K., Hiscock-Anisman, C., & Memon, A. (2009). Training in assessment criteria indicative of deception (ACID) to improve credibility assessments. *Forensic Psychology Practice*, *9*, 199–207. https://doi.org/10.1080/15228930902810078

Colwell, K., Hiscock-Anisman, C., Memon, A., & Michlik, P. M. (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, K., Hiscock-Anisman, C., 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*(3), 167–180.

Colwell, K., Hiscock-Anisman, C. K., & Fede, J. (2013). Assessment criteria indicative of deception (ACID): An example of the new paradigm of differential recall enhancement. In B. Cooper, D. Griesel, & M. Ternes (Eds.), *Applied issues in eyewitness memory, credibility assessment, and investigative interviewing* (pp. 259–291). https://doi.org/10.1007/978-1-4614-5547-9_11

Colwell, K., Hiscock-Anisman, C. K., 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(1), 5–30.

Colwell, K., James-Kangal, N., Hiscock-Anisman, C. K., & Phelan, V. (2015). Should police use ACID: Training and credibility assessments using transcripts versus recordings. *Journal of Forensic Psychology Practice*, 15, 226–247. https://doi.org/10.1080/15228932.2015.1035187

Colwell, K., & Sjerven, E. R. (2005). The "coin-in-hand" stratagem for the forensic assessment of malingering. *American Journal of Forensic Psychology*..

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(4), 293–310. https://doi.org/10.1080/15228932.2011.562804

Colwell, L. H., Colwell, K., Hiscock-Anisman, C. K., 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(1), 68–80. https://doi.org/10.1080/15228932.2012.629592

Dando, C., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police officers to interview tactically. *Journal of Investigative Psychology and Offender Profiling*, 8(2), 189–202. https://doi.org/10.1002/jip.145

Geiselman, R. E. (1984). Enhancement of eyewitness memory: An empirical evaluation of the cognitive interview. *Journal of Police Science and Administration*, 74–80.

Hartwig, M., Granhag, P., Stromwall, L., & Doering, N. (2010). Impression and information management: On the strategic selfregulation of innocent and guilty suspects. *Open Criminology*, 3, 10–16. Retrieved from https://www.researchgate.net/profile/Leif_Stroemwall/publication/229049664_Impression_and_Information_Management_On_the_Strategic_Self-_Regulation_of_Innocent_and_Guilty_Suspects/links/0fcfd50ed2edf2fbd4000000.pdf

Hartwig, M., Granhag, P. A., Strömwall, L., & Andersson, L. (2004). Suspicious minds: Criminals' ability to detect deception. *Psychology, Crime & Law*, 10(1), 83–95. Retrieved from http://0-search.ebscohost.com.www.consuls.org/login.aspx?direct=true&db=aph&AN=11622377&site=ehost-live

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, 13(2), 213–227. https://doi.org/10.1080/10683160600750264

Hartwig, M., Granhag, P. A., 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. https://doi.org/10.1007/s10979-006-9053-9

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, 29(4), 469–484. https://doi.org/10.1007/s10979-005-5521-x

Hauch, V., Sporer, S., Michael, S., & Meissner, C. (2016). Data does training improve the detection of deception? A meta-analysis. *Communication Research*, 43(3), 283–343. https://doi.org/10.1177/0093650214534974

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. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, 11(6), 967–974. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01688638908400949

Liu, M., Granhag, P., Landstrom, S., Hjelmsaeter, E., Strömwall, L., & Vrij, A. (2010). Can you remember what was in your pocket when you were stung by a bee?': Eliciting cues to deception by asking the unanticipated. *Open Criminology*, 3(1), 31–36.

Luke, T. J., Hartwig, M., Brimbal, L., Chan, G., Jordan, S., Joseph, E., … Granhag, P. A. (2012). Interviewing to elicit cues to deception: Improving strategic use of evidence with general-to-specific framing of evidence. *Journal of Police and Criminal Psychology*, 28(1), 54–62.

Mann, S., & Vrij, A. (2006). Police officers' judgements of veracity, tenseness, cognitive load and attempted behavioural control in real-life police interviews. *Psychology, Crime & Law*, 12(3), 307–319.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: police officers' ability to detect suspects' lies. *The Journal of Applied Psychology*, 89(1), 137–149. https://doi.org/10.1037/0021-9010.89.1.137

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, and Law*, 16(4), 340–372. https://doi.org/10.1037/a0020518

Memon, A., Zaragoza, M., Clifford, B., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. *Law and Human Behavior*, 34, 105–117. https://doi.org/10.1007/s10979-008-9172-6

Montalvo, L., Hallinan, C., Hiscock-Anisman, C., Morrissey, D., Bonilla, Y., & Colwell, K. (2013). Using ACID to improve credibility assessments with written and audio statements. *American Journal of Forensic Psychology*, 31(2), 55–69.

Morgan, C., Rabinowitz, Y., Leidy, R., & Coric, V. (2014). Efficacy of combining interview techniques in detecting deception related to bio-threat issues. *Behavioral Sciences & the Law*, 32, 269–285. https://doi.org/10.1002/bsl.2098

Morgan, C. A., Colwell, K., & Hazlett, G. A. (2011). Efficacy of forensic statement analysis in distinguishing truthful from deceptive eyewitness accounts of highly stressful events. *Journal of Forensic Sciences*, 56(5), 1227–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21854383, 1234, DOI: https://doi.org/10.1111/j.1556-4029.2011.01896.x

Porter, S., & Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology*, 15(1), 57–75. Retrieved from http://onlinelibrary.wiley.com/doi/10.1348/135532509X433151/full

Porter, S., Juodis, M., & Leanne, M. (2010). Evaluation of the effectiveness of a brief deception detection training program. *Journal of Forensic Psychiatry and Psychology*, 21(1), 66–76. https://doi.org/10.1080/14789940903174246

Porter, S., & Yuille, J. (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. https://doi.org/10.1007/BF01498980

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

Tekin, S., Granhag, P., Stromwall, L., MacGiolla, E., Vrij, A., & Hartwig, M. (2015). Interviewing strategically to elicit admissions from guilty suspects. *Law and Human Behavior*, 39(3), 244–252. https://doi.org/10.1037/lhb0000131

Tekin, S., Granhag, P., Stromwall, L., & Vrij, A. (2016). How to make perpetrators in denial disclose more information about their crimes. *Psychology, Crime & Law*, 22(6), 561–580. https://doi.org/10.1080/1068316X.2016.1168425

Vallano, J., & Compo, N. S. (2013). A comfortable witness is a good witness: Rapport-building and susceptibility to misinformation in an investigative mock-crime interview. *Applied Cognitive Psychology*, 25(6).

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, 5(1–2), 39–43. https://doi.org/10.1002/jip.82

Vrij, A., Granhag, P., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science*, 20(1), 28–32. https://doi.org/10.1177/0963721410391245

Vrij, A., Leal, S., Mann, S., & Fisher, R. (2012). Imposing cognitive load to elicit cues to deceit: Inducing the reverse order technique naturally. *Psychology, Crime & Law*, 18(6), 579–594. https://doi.org/10.1080/1068316X.2010.515987

Yuille, J. C. (1989). Analysis of the statements of victims, witnesses and suspects. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 175–191). New York: Kluwer Academic/Plenum Publishers.