Chapter 10
The Inconsistent Suspect: A Systematic Review of Different Types of Consistency in Truth Tellers and Liars

Annelies Vredeveldt, Peter J. van Koppen and Pär Anders Granhag

Introduction

"Today you stated that you were home all night on the night in question, but in your first interview with the police, you said that you went out to get cigarettes." This kind of statement is often used in court to cast doubt on the trustworthiness of witnesses and defendants. Pointing out inconsistencies is considered to be one of the most powerful courtroom tactics at a lawyer's disposal (e.g., Aron and Rosner 1985). Its success is likely due to the fact that many people believe that inconsistency is indicative of lying, and conversely, that consistency is a sign of truth telling. These beliefs have been expressed by lay people and legal professionals alike (Brewer et al. 1999; Granhag and Strömwall 2000b; Greuel 1992; Potter and Brewer 1999; Strömwall and Granhag 2003).

The present chapter reviews the available empirical evidence on the diagnostic value of consistency to predict truth telling, and of inconsistency to predict lying. We cover four different types of consistency, namely: (a) within-statement consistency (i.e., the level of consistency between details within the same statement), (b) between-statement consistency (i.e., the level of consistency between statements made by the same suspect), (c) within-group consistency (i.e., the level of consistency between statements made by different suspects), and (d) statement-evidence consistency (i.e., the level of consistency between the suspect's statement and other pieces of evidence). Before reviewing the empirical evidence, we will explain three

A. Vredeveldt (✉)
VU University Amsterdam, Amsterdam, The Netherlands
e-mail: a.vredeveldt@vu.nl

P. J. van Koppen
VU University Amsterdam, Amsterdam, The Netherlands
e-mail: p.j.van.koppen@vu.nl

Maastricht University, Maastricht, The Netherlands

P. A. Granhag
University of Gothenburg, Göteborg, Sweden
e-mail: pag@psy.gu.se

theory-driven interview approaches designed to amplify differences in consistency between liars and truth tellers.

In light of the variability in the use of terms in the literature on deception, allow us to define the terms used throughout this chapter. For the sake of economy of expression, the term "suspect" will be used as an umbrella term for any person who is asked to provide a statement in experiments (i.e., including witnesses). The term "liar" and "lying" will be used for a person who is instructed not to tell the truth, either via a direct instruction to lie, or via an instruction to discuss an "imagined" experience. The term "statement" will be used to refer to the statement or interview as a whole, whereas "detail" will be used to refer to a specific detail or response within the statement.

Interview Approaches

One of the most frequently reported subjective cues to deception is consistency (e.g., Strömwall and Granhag 2003; Strömwall et al. 2003). Furthermore, both legal professionals and lay people seem to agree on how the consistency cue should be used, namely, consistency is seen as a sign of truth telling, whereas inconsistency is considered to be indicative of lying. However, what we know about memory might suggest exactly the opposite pattern. Probably the most influential approach regarding the relationship between consistency and deception is the "repeat versus reconstruct hypothesis" proposed by Granhag and Strömwall (1999, 2000a). The hypothesis consists of two important propositions. First, the authors highlight the French philosopher Montaigne's argument that a liar should have a good memory. Thus, good liars know that they need to keep track of their lies to avoid being unmasked, and will attempt to carefully repeat the same story. Second, research shows that memory is a reconstructive process (e.g., Baddeley et al. 2009; Loftus 2003). Therefore, when truth-telling suspects rely on their memory for an event, their testimony is likely to contain various types of inconsistencies. In short, it is predicted that the "repeat" strategy used by liars will promote consistency, whereas the "reconstructive" strategy used by truth tellers will undermine consistency. As a consequence, consistency may be a sign of lying, rather than truth telling.

A number of interview approaches have been designed to maximize differences between liars and truth tellers. Most of these approaches take advantage of the cognitive differences between liars and truth tellers specified in the repeat versus reconstruct hypothesis. Here, we review three cognitive approaches most commonly discussed in the literature on consistency, namely, asking unanticipated questions, imposing cognitive load, and strategic use of evidence (see Vrij and Granhag 2012; Vrij et al. 2010a, 2011 for more elaborate reviews).

Unanticipated Questions

Liars are often able to appear consistent because they are likely to have prepared for the interview. Indeed, Hartwig et al. (2007) found that liars were significantly more

likely than truth tellers to have a strategy prior to the interview (see also Hartwig et al. 2010), and that one of their most frequent strategies was to "stick with the story". Preparation is particularly important when there is another suspect involved, in which case the suspects need to "get their stories straight" (Vrij et al. 2009). In contrast, truth tellers are less likely to prepare for an interview. Kassin (2005) argues that innocence puts innocents at risk because (a) people tend to believe in a just world (Lerner 1980), in which innocent people do not get convicted, and (b) people believe that their true feelings and intentions will be apparent to the interviewer, a fallacy referred to as the "illusion of transparency" (Gilovich et al. 1998). Due to the reconstructive nature of memory, truth tellers' unprepared stories may thus appear less consistent than liars' prepared stories.

Vrij et al. (2009) point out that investigators can exploit liars' preparation strategies by asking questions that liars did not anticipate prior to the interview. For instance, Vrij et al. asked questions that included spatial shifts (e.g., "In relation to the front door, where did you and your friend sit?") or temporal shifts (e.g., "In which order did you discuss the different topics you mentioned earlier?"). In response to these questions, liars will likely have no prepared answers ready. This increases the likelihood that they will have to think on the spot, thereby showing signs of cognitive load (cf. Lancaster et al. 2012; Liu et al. 2010). Furthermore, during subsequent interviews, liars may not remember the answers they fabricated hastily during the interview as accurately as the answers they had prepared extensively prior to the interview (Leins et al. 2012). Finally, in the case of multiple suspects, liars will be more likely to contradict each other in response to unanticipated questions than in response to anticipated questions (Vrij et al. 2009). Later in the chapter, we will review how asking unanticipated questions during the interview affects different types of consistency in liars and truth tellers.

Cognitive Load

A related interviewing approach is asking questions that impose cognitive load on the suspect (Vrij et al. 2008a). An important assumption of this approach is that lying involves a number of mentally taxing processes, including the creation of the lie, suppression of the truth, and monitoring one's own and others' behavior. Because lying is usually a more cognitively demanding task than telling the truth (Buller and Burgoon 1996; Zuckerman et al. 1981), lying suspects should typically have fewer cognitive resources available during the interview than truth-telling suspects. This assumption has been supported by findings that liars perform worse than truth tellers on secondary tasks during interviews (Lancaster et al. 2012). Furthermore, liars are not always able to suppress subtle signs of cognitive load. For instance, liars typically move their hands and fingers less than truth tellers due to high cognitive load (Caso et al. 2006; Vrij 2006; Vrij et al. 1997, 2008b; but see DePaulo et al. 2003). Nevertheless, these subtle signs are generally not sufficient for observers to detect deception—both trained and untrained lie-catchers tend to perform only slightly above chance level (Bond and DePaulo 2006; DePaulo et al. 2003; Hartwig and Bond 2011).

The investigative interviewer can amplify observable differences between liars and truth tellers by providing instructions that require complex mental operations. For instance, Vrij et al. (2008b) found that suspects who were instructed to tell their story in reverse chronological order showed significantly more signs of cognitive load (e.g., more speech hesitations and increased mention of cognitive operations) than suspects who told their story in normal chronological order. Furthermore, observers were significantly better at distinguishing between lying and truth-telling suspects when judging the reverse-order statements compared to the chronological-order statements. In a similar vein, Vrij et al. (2010b) found that an instruction to maintain eye contact with the interviewer (which imposes cognitive load; cf. Glenberg et al. 1998; Markson & Paterson 2009; Vredeveldt et al. 2011) amplified cues to deception compared to a control condition, and increased observers' ability to distinguish between liars and truth-tellers. In sum, imposing additional cognitive load during the interview has the potential to highlight observable differences between liars and truth tellers.

Strategic Use of Evidence

A third approach to amplifying differences between lying and truth-telling suspects involves the strategic use of evidence. When investigators have independent evidence available (e.g., fingerprints or witnesses), they can make a strategic decision about how to use that evidence in the investigation. Hartwig et al. developed the Strategic Use of Evidence (SUE) technique (Hartwig et al. 2005, 2006), which involves a strategic and a tactical level. The strategic, more abstract level concerns general cognitive principles involved in interviews, such as the tendency for truth tellers to be more forthcoming than liars (see Granhag and Hartwig 2008, for more on the strategic level of the SUE technique).

The tactical level of the SUE technique is case-dependent, and involves concrete interview tools derived from the conceptual framework underlying the SUE technique. SUE tactics can be categorized into three groups: (1) *evidence tactics*, used to evaluate the evidence during preinterview planning; (2) *question tactics*, used to exhaust potential alternative explanations that may account for the evidence; and (3) *disclosure tactics*, used to disclose the evidence in the most effective way. For instance, one tactic that was the focus of the earlier work on the SUE technique (e.g., Hartwig et al. 2005, 2006) is late disclosure of evidence. By disclosing the evidence at the end of the interview rather than at the beginning, the interviewer gives suspects the opportunity to contradict such evidence or their earlier statements. Late disclosure of evidence has proven effective in improving the detection of deception; in one study, police recruits trained in this tactic achieved 85 % deception detection accuracy, compared to 56 % accuracy achieved by their untrained colleagues (Hartwig et al. 2006).

In more recent work, Granhag et al. (2012a, 2012b) have further developed the tactical level of the SUE technique. Based on the Evidence Framing Matrix

(Granhag 2010), they designed a revised SUE technique, referred to as SUE-Incremental (as opposed to SUE-Basic). The Evidence Framing Matrix specifies two important dimensions for each piece of evidence, namely, the strength of the source of the evidence, and the degree of precision of the evidence. In their disclosure of evidence, interviewers can frame the source of the evidence as weak (e.g., "we have information") or strong (e.g., "we have video footage"), and the precision of the evidence as low (e.g., "you were seen in the neighborhood") or high (e.g., "you were seen entering the house"). The SUE-Incremental technique advocates a gradual approach of evidence disclosure, starting with the vaguest phrasing (weak source/low specificity) and gradually moving on to the most precise phrasing (strong source/high specificity). The incremental disclosure of evidence is expected to cause suspects to change their story in order to fit the evidence presented to them, thus eliciting inconsistencies. The SUE-Incremental technique allows for interviewers to make optimal use of the evidence, even when there is only one piece of evidence available in a case.

Another potentially useful addition to interviewers' tactical toolbox is the Tactical Use of Evidence (TUE) approach proposed by Dando et al. (Dando and Bull 2011; Dando et al. 2013). In line with the principles underlying the SUE-Incremental technique, TUE also advocates a "drip-feeding" approach to revealing evidence. However, the empirical evidence on the effectiveness of TUE is currently limited to two studies (Dando and Bull 2011; Dando et al. 2013), and these studies did not assess cues to deception. Hence, it remains to be seen whether TUE can maximize differences in (verbal) behavior between liars and truth tellers.

In sum, the tactical interview tools derived from the conceptual framework underlying strategic use of evidence can improve observers' ability to detect deception. Later in this chapter, we will review how these interview tactics affect different types of consistency.

Consistency and Deception

We now turn to a review of the empirical evidence on the relationship between consistency and deception. The purpose of this review is not only to determine whether there are significant differences in consistency between liars and truth tellers, but also to provide an estimate of the predictive value of consistency cues. To this end, we provide effect sizes for all studies in this review (see Tables 10.1 and 10.2). Because the predictive value of consistency cues depends heavily on the interview technique employed, we report effect sizes separately for the different interview conditions within each study.

In the empirical literature, "consistency" has been operationalized in different ways. One way of assessing consistency is by rating the extent to which a statement as a whole overlaps with another statement (provided at a different time or by a different suspect) or with another piece of evidence (e.g., fingerprints found at the scene of the crime), on a scale ranging from "inconsistent" to "consistent," or vice

versa. For studies reporting such ratings of overall consistency, we calculated the standardized mean difference (SMD) between liars and truth tellers (see Appendix 1 for more information on calculations). An overview of these differences is provided in Table 10.1.

An alternative approach to operationalizing consistency is to count the number of consistent and inconsistent details appearing in statements provided by liars and truth tellers, respectively. For studies reporting this measure, we calculated three measures commonly used to compare effect sizes (see Table 10.2). First, to evaluate the common beliefs that (a) consistency is predictive of truth telling and (b) inconsistency is predictive of lying, we assessed diagnostic values for each of these beliefs. Second, we calculated the odds ratio (OR), which compares the odds of consistency when a suspect is telling the truth with the odds of consistency when a suspect is lying (see Appendix 1 for more information). Third, we provide the natural log of the OR—a measure derived from the OR that is frequently used in meta-analytic comparisons because it is normally distributed and less sensitive to small frequencies than the OR.

Within-Statement Consistency

We define "within-statement consistency" as the correspondence between details provided by a suspect in the space of one statement. Some researchers have examined within-statement consistency in terms of the number of consistent and inconsistent details appearing in the statement (e.g., Walczyk et al. 2009), whereas others have evaluated it in terms of consistency ratings for the statement overall (e.g., Granhag et al. 2012b).

Walczyk et al. (2009) instructed one group of participants to lie and another group to tell the truth in response to questions about their biographical information and recent activities. In their first experiment, they found that the odds of consistency were nearly seven times higher in the truth-telling condition than in the lie condition (see Table 10.2). In their second experiment, Walczyk et al. gave one group of liars the opportunity to rehearse their lies prior to questioning. Rehearsal made the liars somewhat more consistent, but not dramatically so: the odds of consistency for truth tellers were ten times higher than for unrehearsed liars, and still eight times higher than for rehearsed liars. In a follow-up study (Walczyk et al. 2012), participants were asked to answer questions about witnessed videotaped events, instead of their own recent activities. In this study, the odds of consistency were nearly eight times higher for participants who told the truth compared to unrehearsed liars, and 6.5 times higher compared to participants who had rehearsed their lies. In sum, the findings reported by Walczyk et al. suggest that truth tellers are substantially more likely to provide consistent responses within the same interview than liars.

A somewhat more nuanced picture is provided by studies using more realistic methodologies and different interview approaches. For instance, Leins et al. (2011) asked participants to come to the laboratory in pairs. The experimenter sent half of the pairs to lunch together (innocent condition). The remaining pairs did not go to lunch, but were instructed to steal money from a purse in a nearby room (guilty

Table 10.1 Summary statistics for studies reporting consistency ratings

Consistency type	Study	Condition	Truth tellers		Liars		SMD
			M	SD	M	SD	
Within-statement	Leins et al. (2011)		5.6	1.5	3.8	1.6	0.2
	Granhag et al. (2012a)	Early disclosure	4.0	0.0	4.0	0.0	0.0
		SUE-Basic	4.0	0.0	3.6	1.1	0.5
		SUE-Incremental	4.0	0.0	3.0	1.2	1.2
	Granhag et al. (2012b)	Early disclosure	4.0	0.0	3.9	0.6	0.3
		SUE-Basic	4.0	0.0	3.7	0.7	0.7
		SUE-Incremental	4.0	0.0	3.4	0.9	1.0
Between-statement	Granhag et al. (2003)		7.8	1.9	8.7	1.0	−0.1
	Leins et al. (2012)[a]	Same-mode	1.0	0.1	0.9	0.1	0.1
		Different-mode	0.9	0.1	0.7	0.2	0.5
Within-pair	Granhag et al. (2003)	Statement 1	7.5	1.5	8.7	0.7	−0.3
		Statement 2	7.6	1.5	8.1	1.2	−0.1
	Vrij et al. (2009)	Anticipated	4.3	1.4	4.4	1.6	0.0
		Unanticipated[b]	4.4	1.4	3.2	1.4	0.2
Within-triad	Roos af Hjelmsäter et al. (2012)	Anticipated[c]	4.6	1.1	4.4	1.3	0.0
		Unanticipated[d]	4.2	1.4	3.1	1.5	0.2
	Granhag et al. (2012a)	Early disclosure	4.0	0.0	3.6	0.8	0.7
		SUE-Basic	4.0	0.0	2.9	1.2	1.2
		SUE-Incremental	4.0	0.0	2.7	1.4	1.2
Statement-evidence	Hartwig et al. (2005)	Early disclosure	2.8	0.4	2.6	0.6	0.2
		Late disclosure	3.9	0.4	2.2	0.7	1.1
	Jordan et al. (2012)	Early disclosure	5.9	0.3	3.9	2.3	1.0
		Late disclosure	5.2	1.9	1.1	1.8	0.7
	Clemens et al. (2010)	Early disclosure	3.0	0.2	2.5	0.8	0.4
		Late disclosure	2.9	0.3	2.1	0.8	0.5
	Hartwig et al. (2006)	Untrained	1.8	0.3	1.3	0.5	0.4
		SUE	1.9	0.2	0.9	0.5	0.9
	Granhag et al. (2012a)	Early disclosure	4.0	0.0	3.9	0.4	0.5
		SUE-Basic	4.0	0.0	3.3	1.3	0.7
		SUE-Incremental	4.0	0.0	2.5	1.8	1.2
	Granhag et al. (2012b)	Early disclosure	4.3	1.5	1.6	1.4	0.3
		SUE-Basic	4.2	1.3	0.9	1.0	0.4
		SUE-Incremental	5.3	0.7	1.5	1.5	0.9
	Hartwig et al. (2011)	Free recall	2.8	0.2	2.5	0.3	0.4
		Probes	3.0	0.0	2.1	0.5	2.4
		Free recall + probes	2.9	0.3	2.2	0.5	0.7

Means and standard deviations for truth tellers and liars, respectively, and the standard mean difference between truth tellers and liars, by interview condition. *M* mean, *SD* standard deviation, *SMD* standard mean difference

[a] The consistency proportions were treated as scores on a scale from 0 to 1

[b] Averaged over four unanticipated questions

[c] Averaged over four anticipated description questions

[d] Averaged over six unanticipated drawing questions

Table 10.2 Summary statistics for studies reporting frequencies

Consistency type	Study	Consistency	Truth tellers		Liars		DV	OR	Log(OR)
			N		N				
Within-statement	Walczyk et al. (2009, Exp. 1)	Consistent	1339	98.3 %	1,220	89.6 %	1.1	–	–
		Inconsistent	23	1.7 %	142	10.4 %	6.2	–	–
		Odds (consistency)	–	58.2	–	8.6	–	6.8	1.9
	Walczyk et al. (2009, Exp. 2)[a]	Consistent	881	98.9 %	1,615	90.6 %	1.1	–	–
		Inconsistent	10	1.1 %	167	9.4 %	8.4	–	–
		Odds (consistency)	–	88.1	–	9.7	–	9.1	2.2
	Walczyk et al. (2012)[a]	Consistent	884	99.0 %	1,684	93.2 %	1.1	–	–
		Inconsistent	9	1.0 %	122	6.8 %	6.7	–	–
		Odds (consistency)	–	98.2	–	13.8	–	7.1	2.0
	Vrij et al. (2012)[b]	Consistent	29	93.5 %	23	74.2 %	1.3	–	–
		Inconsistent	2	6.5 %	8	25.8 %	4.0	–	–
		Odds (consistency)	–	14.5	–	2.9	–	5.0	1.6
Between-statement	Granhag and Strömwall (2003)	Repeated[c]	144	45.7 %	64	41.6 %	1.1	–	–
		Omissions[c]	171	54.3 %	90	58.4 %	1.1	–	–
		Commissions[d]	81	36.0 %	74	53.6 %	1.5	–	–
		Odds (repeated)	–	0.8	–	0.7	–	1.2	0.2
		Odds (omissions)	–	1.2	–	1.4	–	0.8	-0.2
		Odds (commissions)	–	0.6	–	1.2	–	0.5	-0.7
	Granhag et al. (2003)	Repeated[c]	127	69.8 %	107	62.6 %	1.1	–	–
		Omissions[c]	55	30.2 %	64	37.4 %	1.2	–	–
		Commissions[d]	76	37.4 %	41	27.7 %	0.7	–	–
		Odds (repeated)	–	2.3	–	1.7	–	1.4	0.3
		Odds (omissions)	–	0.4	–	0.6	–	0.7	-0.3
		Odds (commissions)	–	0.6	–	0.4	–	1.6	0.4

Table 10.2 (continued)

Consistency type	Study	Consistency	Truth tellers N	Liars N	DV	OR	Log(OR)
	Strömwall and Granhag (2005)	Repeated[c]	76 33.6%	97 43.1%	0.8	–	–
		Omissions[c]	150 66.4%	128 56.9%	0.9	–	–
		Commissions[d]	106 58.2%	70 41.9%	0.7	–	–
		Odds (repeated)	*0.5*	*0.8*	–	*0.7*	*-0.4*
		Odds (omissions)	*2.0*	*1.3*	–	*1.5*	*0.4*
		Odds (commissions)	*1.4*	*0.7*	–	*1.9*	*0.7*
Within-pair	Wagenaar and Dalderop (1994)[e]	Consistent	39.0%	79.0%	0.5	–	–
		Inconsistent	61.0%	21.0%	0.3	–	–
		Odds (consistency)	*0.6*	*3.8*	–	*0.2*	*-1.8*
	Strömwall and Granhag (2007)	Consistent	97 88.2%	84 76.4%	1.2	–	–
		Inconsistent	13 11.8%	26 23.6%	2.0	–	–
		Odds (consistency)	*7.5*	*3.2*	–	*2.3*	*0.8*
	Vredeveldt and Wagenaar (2013)	Consistent	96 72.7%	50 36.2%	2.0	–	–
		Inconsistent	36 27.3%	88 63.8%	2.3	–	–
		Odds (consistency)	*2.7*	*0.6*	–	*4.7*	*1.5*

Number and percentage of consistent and inconsistent details provided by truth tellers and liars; diagnostic values of consistency to predict truth-telling and of inconsistency (including omissions and commissions) to predict lying; the odds of consistency for truth tellers and liars, respectively; the odds ratio; and the natural log of the odds ratio. *DV* diagnostic values, *OR* odds ratio, *Log(OR)* natural log of the odds ratio

[a] Because there were only minor differences between unrehearsed and rehearsed liars, frequencies were summed across both lie conditions

[b] Vrij et al. counted the number of participants who contradicted themselves rather than the number of responses

[c] The numerator consists of the total number of details reported in the former of the two interviews

[d] The numerator consists of the total number of details reported in the latter of the two interviews

[e] Reflecting proportions of explicit details. Frequencies were not reported

condition). When the latter pairs returned, they were instructed to fabricate an alibi of having had lunch together during the time of the theft. Subsequently, all participants were informed that they were suspected of theft and interviewed individually (either ten minutes or a week later). During the interview, suspects were asked a number of general (anticipated) questions, as well as a number of unanticipated questions, including questions about the spatial layout of the restaurant. Furthermore, they were asked to draw a sketch of the layout of the restaurant. Leins et al. coded the level of consistency between the drawing and the verbal answers about the spatial layout of the restaurant, and found that truth tellers were significantly more consistent than liars (see Table 10.1).

In another recent study, Vrij et al. (2012) sent participants on a mission, during which they were intercepted by one interviewer from a friendly organization and one interviewer from a hostile organization (in counterbalanced order). Each interviewer asked four questions, two of which required participants to describe their route in chronological order, and two of which asked for a description of the route in reverse order. Participants were instructed to tell the truth about their mission to the interviewer from the friendly organization, but to lie to the interviewer from the hostile organization. Vrij et al. found that suspects were substantially more likely to contradict themselves within an interview when they were lying than when they were telling the truth (see Table 10.2).

Granhag et al. (2012b) examined the impact of three interviewing techniques on the levels of within-statement consistency displayed by innocent and guilty suspects. They instructed participants to go into a bookstore, where some participants stole a book (guilty condition), whereas others merely checked the price of the book (innocent condition). Subsequently, participants were accused of theft and interviewed. All participants were instructed to deny stealing the book. During the interview, the evidence was disclosed either before suspects provided their statement (Early Evidence), at the end of the interview (SUE-Basic), or in increments of strength and precision (SUE-Incremental). In the Early Evidence condition, there was no significant difference in within-statement consistency between liars and truth tellers. In contrast, in both SUE conditions, liars were significantly less consistent than truth tellers. The difference was most pronounced in the SUE-Incremental condition (see Table 10.1).

In a follow-up study, Granhag et al. (2012a) examined within-statement consistency in groups of three suspects. All groups were instructed to go into a room, where half of the groups checked some reference numbers and stole something from one of the packages, whereas the other groups only checked the reference numbers. Subsequently, all participants were accused of theft and interviewed individually, using one of the three interview techniques discussed above. In the Early Evidence condition, ratings of within-statement consistency were at ceiling for both truth-telling and lying suspects. When interviewed with SUE-Basic, lying suspects were slightly less consistent than truth-telling suspects, though the difference was not significant. When interviewed with SUE-Incremental, however, lying suspects were significantly less consistent than truth-telling suspects (see Table 10.1).

In sum, it seems that interview techniques may be used to improve the predictive value of within-statement consistency cues. Studies using artificial experimental set-ups, in which the "interview" task is automated (Walczyk et al. 2009, 2012),

typically find that truth tellers are substantially more consistent than liars. In more naturalistic settings, on the other hand, in which suspects are interviewed by a trained interviewer (Granhag et al., 2012a, b), there seem to be smaller differences in within-statement consistency between liars and truth tellers. However, these differences can be amplified by asking unanticipated questions that are likely to impose additional cognitive load (Leins et al. 2011; Vrij et al. 2012), and by strategic disclosure of evidence, preferably in an incremental fashion (Granhag et al. 2012a, b).

Between-Statement Consistency

Another approach is to evaluate the consistency between two consecutive statements provided by the same suspect. Between-statement consistency can be evaluated not only in terms of the number of contradictions between statements, but also in terms of the extent to which two statements overlap. When looking at repeated statements over time, the degree of overlap of two statements is typically broken down into measures of "repetitions" (i.e., details that are mentioned during all statements), "omissions" (i.e., details that are mentioned in an earlier statement but not in a later statement), and "commissions" (i.e., details that are mentioned in a later statement but not in an earlier statement). For instance, Granhag and Strömwall's (2002) participants viewed a staged violent event and were interviewed on three occasions over a period of 11 days (see also Granhag and Strömwall 2001). Half of the participants were instructed to tell the truth about the witnessed event, whereas the other half were instructed to lie in order to cover up for the perpetrators. For 50 % of truthful and 50 % of deceptive participants, the frequencies of repetitions, omissions, commissions, and contradictions between the three statements were coded for four forensically relevant categories of information (e.g., which factor triggered the conflict). There were no direct contradictions in any of the statements, and there was no significant difference between liars and truth tellers in the proportion of details repeated or omitted (see Table 10.2). However, liars were significantly more likely than truth tellers to add details to later statements that they had not mentioned in earlier statements (i.e., commissions; see Table 10.2).

In a follow-up study conducted by Granhag et al. (2003), participants arrived at the laboratory in pairs. Half of the pairs had lunch together and were subsequently interviewed about the lunch (truth tellers). The remaining pairs did not have lunch together, but were instructed to state that they had had lunch together (liars). Granhag et al. included several measures of between-statement consistency. First, they assessed the number of repetitions, omissions, commissions, and contradictions between the two consecutive statements. There were only three contradictions in total, which were not analyzed further. There were no significant differences between liars and truth tellers in the number of repetitions and omissions, but this time it was the truth tellers who included significantly more new information than liars (see Table 10.2). In addition, Granhag et al. asked 120 independent participants to rate the consistency of statements on a scale from 1 (very low degree) to 10 (very high degree; see also Strömwall et al. 2003). Liars' statements tended to be rated as more

consistent than truth tellers' statements, but this difference was not significant (see Table 10.1).

Strömwall and Granhag (2005) investigated the consistency between repeated statements made by 11-year old children. The children participated in two interviews, 1 week apart, about a magician show that they had either experienced (truth tellers) or imagined ("liars"). Strömwall and Granhag coded the number of repetitions, omissions, commissions, and contradictions in the children's statements, and found that the statements about real events contained significantly more omissions and commissions than the statements about imagined events, and slightly fewer repetitions (see Table 10.2). In other words, lying children were more consistent than truth-telling children.

Taken together, the findings reported by Granhag, Strömwall, and colleagues suggest that differences in between-statement consistency between liars and truth tellers are modest at best. When differences exist, the evidence suggests that consistency is indicative of lying rather than truth telling. This finding is in line with the repeat versus reconstruct hypothesis (Granhag and Strömwall 1999, 2000a), but at odds with popular beliefs about the relationship between consistency and lying (e.g., Potter and Brewer 1999). One exception to this general pattern of findings was that Granhag and Strömwall (2002) found that liars were significantly more likely to add new details to later statements than truth tellers. The reason for this exception is unclear, but it may have been due to methodological differences between the studies. For instance, Granhag and Strömwall's (2002) truth-telling participants were given extensive time to prepare their story, whereas truth-telling participants in the other studies (Granhag et al. 2003; Strömwall and Granhag 2005) did not receive time to prepare. Therefore, truth tellers in the latter studies may have had to "reconstruct" more than truth tellers in Granhag and Strömwall's (2002) study, resulting in less consistent testimony.

To the authors' knowledge, only one study has assessed between-statement consistency using a nonstandard interview approach. Leins et al. (2012) instructed half of their participants to perform a number of tasks (e.g., turn on the radio, complete a puzzle). The remaining participants did not perform the tasks but were instructed to convince the interviewer that they had. Participants were interviewed twice, reporting about the items in the room either verbally or by drawing a sketch. There were four interview conditions: verbal–verbal, pictorial–pictorial, verbal–pictorial, or pictorial–verbal. Overall, consistency between the two interviews was significantly higher for truth tellers than for liars. Furthermore, the difference between liars and truth tellers was larger when participants were asked to report in different response modes (verbal–pictorial or pictorial–verbal) than when they were asked to report in the same response modes (verbal–verbal or pictorial–pictorial; see Table 10.1). Leins et al. explain this finding in terms of cognitive flexibility. Truth tellers experienced the event perceptually, allowing them to draw on a rich memory trace when reporting the event in different ways. In contrast, liars likely only had a conceptual representation of the imagined events, reducing their flexibility in reporting the event in various modalities. In sum, it seems that differences in between-statement consistency between liars and truth tellers may be amplified by asking suspects to provide reports in varied response modes.

Within-Group Consistency

When there are multiple suspects in a case, it becomes possible to evaluate the degree of consistency between statements made by different suspects. Wagenaar and Dalderop (1994) were the first to assess this type of consistency. In their study, six pairs of participants went to the zoo, and six other pairs were asked to invent a story about going to the zoo together. Subsequently, all participants were interviewed individually about the zoo visit. In line with the repeat versus reconstruct hypothesis, Wagenaar and Dalderop found that lying pairs were significantly more consistent than truth-telling pairs (see Table 10.2). In a similar vein, in the previously discussed study involving pairs of suspects providing statements about a lunch meeting, Granhag et al. (2003) assessed the extent to which the two suspects' statements overlapped. The statements of lying suspects were found to contain significantly more overlapping themes than the statements of truth-telling suspects (see Table 10.1).

Two studies to date have examined the impact of nonstandard interview approaches on within-group consistency in adult suspects. First, Vrij et al. (2009) used a similar methodology as Granhag et al. (2003), in which pairs of suspects either had lunch together or invented a story about having lunch. However, Vrij et al. introduced a number of unanticipated questions during the interview (e.g., "Who finished his food first?") in addition to the more standard, anticipated questions (e.g., "What did you do in the restaurant?"). For the anticipated questions, there was no significant difference in correspondence between lying and truth-telling pairs. In response to the unanticipated questions, however, lying pairs of suspects contradicted each other significantly more than truth-telling pairs (see Table 10.1). Second, Granhag et al. (2012a) extended previous research on within-group consistency by assessing consistency for groups of three suspects (i.e., within-triad consistency), and by evaluating the impact of three interview techniques (Early Evidence, SUE-Basic, and SUE-Incremental). They used a mock theft paradigm, after which suspects were interviewed individually. Overall, within-triad consistency was significantly higher for truth-telling groups of suspects than for lying groups. Furthermore, the difference between lying and truth-telling triads was significant for the SUE-Basic and the SUE-Incremental conditions, but not for the Early Evidence condition (see Table 10.1).

In addition, three studies to date have investigated within-group consistency for younger participants. Strömwall and Granhag (2007) recruited pairs of adolescents (aged 12–13), half of whom experienced an encounter with an unknown man, whereas the other half were instructed to imagine and discuss such an encounter. All adolescents were interviewed individually. Truth-telling pairs of adolescents were found to be significantly more consistent than lying pairs (see Table 10.2). Vredeveldt and Wagenaar (2013) conducted a similar study with pairs of children aged 8–10. Again, it was found that truth-telling pairs provided significantly more consistent responses than lying pairs (see Table 10.2)—a finding at odds with Granhag et al.'s (2003) findings for adult pairs. Finally, in Roos af Hjelmsäter et al.'s (2012) study, adolescents aged 13–14 participated in groups of three. They experienced or imagined the same event as Strömwall and Granhag's (2007) participants, after which

they were interviewed individually. In addition to the anticipated questions used by Strömwall and Granhag, adolescents in Roos af Hjelmsäter et al.'s study were given an unanticipated task, namely, drawing the positions of the actors on a spatial layout. Subsequently, adult observers rated the truth-telling triads as significantly more consistent than the lying triads. Interestingly, the difference between liars and truth tellers was only significant for the unanticipated task.

In sum, in "standard" interview settings, lying groups of adult suspects are often found to be either more consistent (Granhag et al. 2003; Wagenaar and Dalderop 1994) or equally consistent (Granhag et al. 2012a; Vrij et al. 2009) as truth-telling groups of suspects. In contrast, lying groups of children or adolescents are typically found to be *less* consistent than their truth-telling counterparts (Roos af Hjelmsäter et al. 2012; Strömwall and Granhag 2007; Vredeveldt and Wagenaar 2013). The difference between adults and children might be explained in light of developmental differences in social and cognitive functioning. Children are typically less aware of their own and other's mental states than adults are (Gallup 1998; Johnson et al. 2005), which makes them less likely to grasp the importance of appearing consistent with their group members in order to convince others of their truthfulness. Furthermore, even if children are aware of the importance of appearing consistent, they may be less skilful at controlling the verbal content of their statements than adults are (cf. Talwar & Lee 2002). Finally, children typically behave more egocentrically than adults (Epley et al. 2004), and hence may be less successful in collaborating with their peers to create a mutually coherent story.

Regardless of age, certain interview techniques have been found to increase differences between lying and truth-telling groups of suspects. Thus, when the interviewer poses unanticipated questions (Roos af Hjelmsäter et al. 2012; Vrij et al. 2009) or uses evidence strategically (Granhag et al. 2012a), more inconsistencies between lying suspects may be exposed.

Statement-Evidence Consistency

In cases in which police interviewers have access to evidence that is independent from the statements provided by suspects, they have the opportunity to examine an additional index of consistency, namely the extent to which the statement is consistent with the other evidence.

Hartwig et al. (2005) were among the first to examine consistency in a scenario in which other evidence was available. They used a mock crime paradigm, in which participants in the guilty condition were instructed to steal a wallet from a briefcase in a video store, whereas participants in the innocent condition only looked for a particular movie (moving the briefcase during their search) and left the store empty-handed. After a week, all participants came back to the laboratory and were informed that they would be interviewed due to suspicions of theft. During the interview, they were presented with three pieces of evidence (e.g., fingerprints on the briefcase), which could be viewed as incriminating, but could also be explained by the innocent

task alibi. The evidence was disclosed either early or late in the interview. To measure statement-evidence consistency, Hartwig et al. coded the extent to which suspects contradicted the evidence when they were asked questions about the event. In the early-disclosure condition, liars and truth tellers were found to be equally consistent with the evidence. In the late-disclosure condition, however, liars were significantly less consistent than truth tellers (see Table 10.1).

Jordan et al. (2012) used the same experimental paradigm as Hartwig et al. (2005) and found that truthful statements were significantly more consistent with the evidence than deceptive statements in both the early- and late-disclosure conditions. However, in line with Hartwig and colleagues' findings, the absolute difference between liars and truth tellers was larger in the late-disclosure condition than in the early-disclosure condition (see Table 10.1).[1] In a similar mock crime paradigm, Clemens et al. (2010) examined the effect of late disclosure of evidence on statement-evidence consistency in statements provided by adolescents aged 12–14. Overall, truth-telling adolescents were more consistent with the evidence than lying adolescents (see Table 10.1). Furthermore, for the most incriminating piece of evidence, liars in the late disclosure condition were significantly less consistent with the evidence than liars in the early-disclosure condition.

Hartwig et al. (2006) trained 41 police trainees in the use of the SUE technique, whereas 41 other police trainees were not trained. These trainees then interviewed participants suspected of a theft, in a similar mock theft paradigm as described above. Statement-evidence consistency was significantly higher for truth tellers than for liars. Moreover, lying suspects interviewed by the SUE trainees displayed significantly lower statement-evidence consistency than lying suspects interviewed by the other trainees, whereas SUE training had no significant impact on the degree of statement-evidence consistency displayed by truth tellers (see Table 10.1). In short, the use of SUE amplified differences in statement-evidence consistency between liars and truth tellers.

Granhag et al. (2012b) also examined the impact of interviewing techniques (Early Evidence, SUE-Basic, and SUE-Incremental) on the degree of statement-evidence consistency displayed by innocent and guilty suspects accused of stealing a book. Across all interview conditions, truth tellers displayed significantly higher statement-evidence consistency than liars, but the size of the difference between liars and truth tellers was largest in the SUE-Incremental condition (see Table 10.1).[2] In a similar vein, Granhag et al. (2012a) examined statement-evidence consistency for guilty and innocent suspects interviewed with one of the same three interview techniques. Statement-evidence consistency tended to be higher for truth

---

[1] Note that the SMD was larger for the early-disclosure condition due to the small standard deviation for innocent suspects in that condition (the accuracy of which was confirmed with the authors of the article).

[2] It should be noted that the large difference between liars and truth tellers in the SUE-Incremental condition was due to the relatively high level of statement-evidence consistency observed for truth tellers in this condition. Contrary to expectations, deceptive statements in the SUE-Incremental condition were nearly as consistent with the evidence as those in the Early Evidence condition.

tellers than for liars in all conditions, but the difference was only significant in the SUE-Incremental condition. Table 10.1 shows that liars in this condition displayed substantially lower statement-evidence consistency than truth tellers.

In most of the studies on statement-evidence consistency, interviewers used a combination of open- and closed-ended questions. Hartwig et al. (2011) examined which of these types of questions is most effective in differentiating between liars and truth tellers based on statement-evidence consistency. They hypothesized that free-recall questions would allow lying suspects to evade mentioning any incriminating evidence, whereas closed-ended probing questions would cause them to contradict the evidence. In a mock-theft paradigm, suspects were assigned to one of three interview conditions: (a) free-recall questions only, (b) closed-ended questions only, or (c) a combination of free recall and closed-ended questions. Overall, deceptive statements were found to be significantly less consistent with the evidence than truthful statements. Furthermore, there was a significant interaction between veracity and interview condition: the difference in statement-evidence consistency between liars and truth tellers was smallest when suspects were only asked free-recall questions, and largest when they were only asked closed-ended probes (see Table 10.1). Thus, closed-ended questions allow for better differentiation between liars and truth tellers based on statement-evidence consistency than free-recall questions.

In sum, empirical findings suggest that truthful statements are typically either equally or more consistent with the other available evidence than deceptive statements are. Furthermore, differences in statement-evidence consistency between liars and truth tellers can be amplified by strategic use of the available evidence (e.g., Hartwig et al. 2006; Hartwig et al. 2005; Jordan et al. 2012). In this respect, incremental disclosure of evidence seems to be even more effective than late disclosure (Granhag et al. 2012a, 2012b), and closed-ended questions seem to be more effective than free-recall questions (Hartwig et al. 2011).

Conclusions and Future Directions

Consistency and Deception

The literature on the relationship between consistency and deception paints a relatively coherent picture. In terms of within-statement, between-statement, and within-group consistency, most studies employing standard face-to-face interviews have found that adult suspects who are lying are typically either equally consistent or more consistent than their truth-telling counterparts. This is in line with the repeat versus reconstruct hypothesis, which suggests that the "repeat" strategy employed by liars results in more consistent statements than truth tellers' "reconstruct from memory" strategy. However, in terms of statement-evidence consistency, liars' statements are typically either equally consistent or *less* consistent with the evidence than truth tellers' statements. Despite the differences depending on the type of consistency under investigation, the accumulative evidence seems to suggest that there is little

empirical basis for the commonly held belief that truth-telling suspects are always more consistent than lying suspects.

Inspection of the effect sizes in Tables 10.1 and 10.2 confirms that differences in consistency between liars and truth tellers tend to be modest. Under "standard" conditions (in which the questions were anticipated or the evidence was disclosed early), the standardized mean differences (displayed in Table 10.1) ranged from negative to positive, and most were small (.20) to medium (.50) in size (as defined by Cohen 1992). In terms of diagnostic values (displayed in Table 10.2), consistency was not very diagnostic of truth telling in any of the studies included in this review, with DVs ranging from 0.5 (Wagenaar and Dalderop 1994) to 2.0 (Vredeveldt and Wagenaar 2013). In contrast, inconsistency was somewhat diagnostic of lying in some of the studies. It should be noted, however, that all studies with relatively high DVs for inconsistency (between 4.0 and 8.4) either employed artificial methodologies (Walczyk et al. 2009, 2012), or imposed cognitive load on the suspects (Vrij et al. 2012). Furthermore, even the highest DV in Table 10.2 does not seem very impressive in light of findings that properly conducted eyewitness identifications have DVs of around 15 (De Jong et al. 2005; Wagenaar and Van Der Schrier 1996). In sum, the popular belief that consistency is predictive of truth telling has not received empirical support, and the belief that inconsistency is predictive of lying has to date only received modest support, in experiments that were rather unrealistic.

Interview Approaches

In light of the relative lack of consistency differences between liars and truth tellers, several interview approaches have been developed to try to improve differentiation between liars and truth tellers. First, asking unanticipated questions during the interview has been found to elicit inconsistencies in liars' statements, in terms of within-statement consistency (Leins et al. 2011), between-statement consistency (Leins et al. 2012), and within-group consistency (Roos af Hjelmsäter et al. 2012; Vrij et al. 2009). Even though the unanticipated-question technique significantly increased differences in consistency between deceptive and truthful statements, the differences between lying and truth-telling suspects who were asked unanticipated questions were still only small to medium in size (see Table 10.1; Leins et al. 2012; Roos af Hjelmsäter et al. 2012; Vrij et al. 2009).

The strategic use of evidence during interviews has also been found to amplify differences between liars and truth tellers in terms of within-statement consistency (Granhag et al. 2012a, b), within-group consistency (Granhag et al. 2012a), and statement-evidence consistency (Granhag et al. 2012a, 2012b; Hartwig et al. 2006; Hartwig et al. 2005; Jordan et al. 2012). Moreover, when evidence was disclosed late in the interview, SMDs between liars and truth tellers were increased to medium (.50) to large (.80) in size, and when evidence was disclosed incrementally, all SMDs were above .80 (see Table 10.1). Thus, based on the evidence to date, strategic use of

evidence seems to be the most effective way of increasing differences between liars and truth tellers.

One important unanswered question is the impact of imposing cognitive load during the interview on various types of consistency. We know from previous research that imposing cognitive load can amplify other cues to deception, such as increases in speech hesitations and decreases in spatial and auditory details in liars' statements (Vrij et al. 2008b; Vrij et al. 2010b). However, none of these studies have assessed consistency. Although Vrij et al. (2012) examined the impact of imposing cognitive load on consistency, they only analyzed contradictions that emerged *between* the chronological-order and reverse-order responses. Future investigations should compare consistency within statements provided under conditions of high cognitive load to consistency within statements provided in control conditions. Interestingly, Vrij et al. (2008b) state that they "have been told by several American investigators who used the reverse-order instruction when interviewing suspects, that suspects frequently gave themselves away with obviously non-credible stories that were replete with inconsistencies" (p. 263). This anecdotal evidence should be assessed more systematically in future research.

Limitations

These recent developments in investigative interviewing approaches provide some promising prospects for deception detection in the real world. Nevertheless, it should be noted that many of the studies on which these approaches are based have limited ecological validity (cf. Van Koppen 2012). First, most studies recruit undergraduate students as participants, a group that is unlikely to be representative of real criminal suspects. In addition, the "crimes" committed by these participants are trivial, and do not have any real consequences for the participants. In fact, in the typical experimental setup, the "liar" is asked to lie by the experimenter. Thus, lying is desired in these experiments, whereas it is usually considered wrong in real police interviews. The conditions for truth tellers may be exactly the other way around. In real police interviews, truth telling is considered desirable and relatively easy to do, whereas truth telling in experiments may sometimes be more demanding than lying. Truth tellers can make mistakes, whereas liars can "win the game" by beating the experimental interviewer. In that sense, the results of research to date may, with a little exaggeration, reflect the psychology of a game played by relatively intelligent students rather than what really goes on in the mind of the average suspect during police interviews.

Although ethical considerations prohibit researchers from alleviating some of these concerns about ecological validity, future research should make every attempt to approach real-world conditions as much as possible. For example, several researchers have evaluated the verbal and nonverbal behavior of individuals pleading for the return of a missing relative, half of whom were subsequently convicted of murdering the relative themselves (Ten Brinke and Porter 2012; Ten Brinke et al. 2012; Vrij

and Mann 2001). This line of research is a step in the right direction toward more realistic deception detection research, and consistencies in such statements could be examined.

Another problem with the research on deception detection pertains to the issue of base rates. Typically, studies in this area employ base rates of 50 % guilty participants and 50 % innocent participants. In real life, however, suspects at the police station are typically there for a reason; that is, they are often suspects because there is strong or reasonable evidence against them. Although the true base rate of guilty suspects can never be known, legal-psychological scholars have estimated that the guilty base rate might be closer to 95 % than 50 % (Clark 2012; Crombag et al. 1992; Van Koppen 2012; Wagenaar 2005). The effectiveness of interview methods is heavily dependent on guilty base rates: if 95 % of suspects are guilty, investigators will achieve higher deception detection accuracy by using methods that are biased towards guilt than by using methods that are biased towards innocence. In fact, if a police officer uses any random method, valid or not, to assess the veracity of denying suspects, she is usually right if the method brings her to the conclusion that the suspect is lying. Of course, implementation of methods that are biased towards guilt is probably not desirable because society may place greater value on protecting the innocent than on convicting the guilty (cf. Clark 2012). Nevertheless, researchers should be more aware of the potentially skewed guilty base rate in the real world.

Conclusion

Taken together, the corpus of research on consistency and deception suggests that—contrary to beliefs held by lay people and legal professionals—consistency is not necessarily indicative of truth telling, and inconsistency is not necessarily a sign of lying. The research on interview approaches designed to amplify cues to deception has been immensely valuable in informing legal professionals about more effective methods of eliciting cues to deception. Nevertheless, the general literature on suspect interviewing to date has perhaps focused a little too much on the detection of deception and the elicitation of confessions.

Looking ahead, we would welcome more research exploring interviewing approaches that promote the elicitation of forensically relevant information from suspects. In the literature on witness interviewing, this issue has received ample attention (see e.g., Memon et al. 2010; Vredeveldt and Penrod 2012; Wagstaff et al. 2011), and Fisher and Perez (2007) have suggested that some interview methods designed for witnesses may also be effective for use with suspects (as recommended since 1992 by the PEACE approach used nationally in England and Wales; Milne and Bull 1999). Eventually, we hope that converging evidence from empirical research will inform the implementation of suspect interview techniques that improve deception detection, increase the diagnosticity of confessions, *and* provide important new leads for further police investigation.

Appendix 1: Effect Size Estimates

Standardized Mean Difference

For studies in which mean consistency scores were obtained, we examined the standardized mean difference (SMD) between liars and truth tellers as an indicator of effect size. Because this measure tends to be upwardly biased when based on small samples, we provide an unbiased estimate of the SMD throughout this chapter (using the correction provided by Hedges (1981), which is depicted in the second part of the equation below), calculated as:

$$\text{SMD} = \left(\frac{\bar{X}_T - \bar{X}_L}{s_p} \right) \left(1 - \frac{3}{4N - 9} \right),$$

where $\bar{X}_T$ is the mean consistency score for truth tellers, $\bar{X}_L$ is the mean consistency score for liars, N is the total sample size and s_p is the pooled standard deviation, calculated as:

$$s_p = \sqrt{\frac{(n_T - 1)s_T^2 + (n_L - 1)s_L^2}{(n_T - 1) + (n_L - 1)}},$$

where n_T is the number of truth tellers, n_L is the number of liars, s_T is the standard deviation for truth tellers, and s_L is the standard deviation for liars.

Odds Ratio

For studies that used frequencies of consistent and inconsistent details for liars and truth tellers as the dependent measure, we examined the odds ratio (OR) as an indicator of effect size. The odds ratio is calculated as:

$$\text{Odds ratio} = \frac{ad}{bc},$$

where a is the number of consistent details provided by truth tellers, b the number of inconsistent details provided by truth tellers, c the number of consistent details provided by liars, and d the number of inconsistent details provided by liars (cf. Lipsey and Wilson 2001). An OR of 1 would indicate no relationship between consistency and truth telling, an OR greater than 1 suggests that consistency is predictive of truth telling, and an OR between 0 and 1 suggests that consistency is predictive of lying. For example, an OR of 4 would indicate that the odds of consistency are four times greater for truth tellers than for liars, whereas an OR of 0.25 would indicate that the odds of consistency are four times *smaller* for truth tellers than for liars.

References

Aron, R., & Rosner, J. L. (1985). *How to prepare witnesses for trial*. Colorado Springs: Shepard's/McGraw-Hill.

Baddeley, A. D., Eysenck, M., & Anderson, M. C. (2009). *Memory*. Hove: Psychology Press.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality & Social Psychology Review, 10,* 214–234. doi:10.1207/s15327957pspr1003_2.

Brewer, N., Potter, R., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology, 13,* 297–313. doi:10.1002/(sici)1099-0720(199908)13:4<297::aid-acp578>3.0.co;2-s.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242. doi:10.1111/j.1468-2885.1996.tb00127.x.

Caso, L., Vrij, A., Mann, S., & De Leo, G. (2006). Deceptive responses: The impact of verbal and non-verbal countermeasures. *Legal and Criminological Psychology, 11,* 99–111. doi:10.1348/135532505x49936.

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform: Psychological science and public policy. *Perspectives on Psychological Science, 7,* 238–259. doi:10.1177/1745691612439584.

Clemens, F., Granhag, P. A., Strömwall, L. A., Vrij, A., Landström, S., Roos af Hjelmsäter, E., & Hartwig, M. (2010). Skulking around the dinosaur: Eliciting cues to children's deception via strategic disclosure of evidence. *Applied Cognitive Psychology, 24,* 925–940. doi:10.1002/acp.1597.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159. doi:10.1037/0033-2909.112.1.155.

Crombag, H. F. M., Van Koppen, P. J., & Wagenaar, W. A. (1992). *Dubieuze zaken: De psychologie van strafrechtelijk bewijs* (4th ed.). Amsterdam: Uitgeverij Contact.

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

Dando, C. J., Bull, R., Ormerod, T. C., & Sandham, A. L. (2013). Helping to sort the liars from the truth-tellers: The gradual revelation of information during investigative interviews. *Legal and Criminological Psychology*, Advance online publication. doi: 10.1111/lcrp.12016.

De Jong, M., Wagenaar, W. A., Wolters, G., & Verstijnen, I. M. (2005). Familiar face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 11,* 87–97. doi:10.1080/10683160410001715123.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129,* 74–118. doi:10.1037/0033-2909.129.1.74.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40,* 760–768. doi:10.1016/j.jesp.2004.02.002.

Fisher, R. P., & Perez, V. (2007). Memory-enhancing techniques for interviewing crime suspects. In S. A. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 329–354). Chichester, UK: John Wiley & Sons, Ltd.

Gallup, G. G. (1998). Self-awareness and the evolution of social intelligence. *Behavioural Processes, 42,* 239–247. doi:10.1016/s0376-6357(97)00079-x.

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology, 75,* 332–346. doi:10.1037/0022-3514.75.2.332.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26,* 651–658. doi:10.3758/BF03211385.

Granhag, P. A. (2010). *The Strategic Use of Evidence (SUE) technique: A scientific perspective.* Paper presented at the High Value Detainee Interrogation Group (HIG, FBI). HIG Research Symposium: Interrogation in the European Union, Washington, D.C.

Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law, 14,* 189–200. doi:10.1080/10683160701645181.

Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations—Stretching the deception detection paradigm. *Expert Evidence, 7,* 163–174. doi:10.1023/a:1008993326434.

Granhag, P. A., & Strömwall, L. A. (2000a). Deception detection: Examining the consistency heuristic. In C. M. Breur, M. M. Kommer, J. F. Nijboer & J. M. Reintjes (Eds.), *New trends in criminal investigation and evidence* (Vol. 2, pp. 309–321). Antwerpen: Intresentia.

Granhag, P. A., & Strömwall, L. A. (2000b). Effects of preconceptions on deception detection and new answers to why lie-catchers often fail. *Psychology, Crime & Law, 6,* 197–218. doi:10.1080/10683160008409804.

Granhag, P. A., & Strömwall, L. A. (2001). Deception detection based on repeated interrogations. *Legal and Criminological Psychology, 6,* 85–101. doi:10.1348/135532501168217.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16,* 243–257. doi:10.1002/acp.784.

Granhag, P. A., Strömwall, L. A., & Jonsson, A.C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology, 33,* 848–868. doi:10.1111/j.1559-1816.2003.tb01928.x.

Granhag, P. A., Rangmar, J., & Strömwall, L. A. (2012a). Small cells of suspects: Eliciting cues to deception by strategic interviewing. *Manuscript submitted for publication.*

Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2012b). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*. doi:10.1111/j.2044-8333.2012.02047.x.

Greuel, L. (1992). Police officers' beliefs about cues associated with deception in rape cases. In F. Lösel, D. Bender & T. Bliesener (Eds.), *Psychology and law —International perspectives* (pp. 234–239). Berlin: Walter de Gruyter.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137,* 643–659. doi:10.1037/a0023589.

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

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews. *Law and Human Behavior, 30,* 603–619. doi:10.1007/s10979-006-9053-9.

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law, 13,* 213–227.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal, 3,* 10–16. doi:10.2174/1874917801003010010.

Hartwig, M., Granhag, P. A., Strömwall, L., Wolf, A. G., Vrij, A., & Roos af Hjelmsäter, E. (2011). Detecting deception in suspects: Verbal cues as a function of interview strategy. *Psychology, Crime & Law, 17,* 643–656. doi:10.1080/10683160903446982.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Johnson, A. K., Barnacz, A., Yokkaichi, T., Rubio, J., Racioppi, C., Shackelford, T. K., Fisher, M. L., Keenan, J. P. (2005). Me, myself, and lie: The role of self-awareness in deception. *Personality and Individual Differences, 38,* 1847–1853. doi:10.1016/j.paid.2004.11.013.

Jordan, S., Hartwig, M., Wallace, B., Dawson, E., & Xhihani, A. (2012). Early versus late disclosure of evidence: Effects on verbal cues to deception, confessions, and lie catchers' accuracy. *Journal of Investigative Psychology and Offender Profiling, 9,* 1–12. doi:10.1002/jip.1350.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60,* 215–228. doi:10.1037/0003-066X.60.3.215.

Lancaster, G. L. J., Vrij, A., Hope, L., & Waller, B. (2012). Sorting the liars from the truth tellers: The benefits of asking unanticipated questions on lie detection. *Applied Cognitive Psychology*. doi:10.1002/acp.2879.

Leins, D. A., Fisher, R. P., Vrij, A., Leal, S., & Mann, S. (2011). Using sketch drawing to induce inconsistency in liars. *Legal and Criminological Psychology, 16,* 253–265. doi:10.1348/135532510x501775.

Leins, D. A., Fisher, R. P., & Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: Detecting deception through varying report modes. *Applied Cognitive Psychology, 26,* 601–607. doi:10.1002/acp.2837.

Lerner, M. J. (1980). *The belief in a just world*. New York: Plenum.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. London: Sage Publications.

Liu, M., Granhag, P. A., Landström, S., Roos af Hjelmsäter, E., Strömwall, L. A., & 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. *The Open Criminology Journal, 3,* 31–36. doi:10.2174/1874917801003010031.

Loftus, E. F. (2003). Our changeable memories: Legal and practical implications. *Nature Reviews Neuroscience, 4,* 231–234. doi:10.1038/nrn1054.

Markson, L., & Paterson, K. B. (2009). Effects of gaze-aversion on visual-spatial imagination. *British Journal of Psychology, 100*, 553–563. doi: 10.1348/000712608X371762.

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,* 340–372. doi:10.1037/a0020518.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester: Wiley.

Potter, R., & Brewer, N. (1999). Perceptions of witness behaviour-accuracy relationships held by police, lawyers and mock-jurors. *Psychiatry, Psychology and Law, 6,* 97–103. doi:10.1080/13218719909524952.

Roos af Hjelmsäter, E., Öhman, L., Granhag, P. A., & Vrij, A. (2012). 'Mapping' deception in adolescents: Eliciting cues to deceit through an unanticipated spatial drawing task. *Legal and Criminological Psychology*. doi:10.1111/j.2044-8333.2012.02068.x.

Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law, 9,* 19–36. doi:10.1080/10683160308138.

Strömwall, L. A., & Granhag, P. A. (2005). Children's repeated lies and truths: Effects on adults' judgments and reality monitoring scores. *Psychiatry, Psychology and Law, 12,* 345–356. doi:10.1375/pplt.12.2.345.

Strömwall, L. A., & Granhag, P. A. (2007). Detecting deceit in pairs of children. *Journal of Applied Social Psychology, 37,* 1285–1304. doi:10.1111/j.1559-1816.2007.00213.x.

Strömwall, L. A., & Granhag, P. A., & Jonsson, A.-C. (2003). Deception among pairs: "Let's say we had lunch and hope they will swallow it!". *Psychology, Crime, and Law, 9*, 109–124. doi:10.1080/1068316031000116238.

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26,* 436–444. doi:10.1080/01650250143000373.

Ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioural consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior, 36,* 469–477. doi:10.1037/h0093929.

Ten Brinke, L., Porter, S., & Baker, A. (2012). Darwin the detective: Observable facial muscle contractions reveal emotional high-stakes lies. *Evolution and Human Behavior, 33,* 411–416. doi:10.1016/j.evolhumbehav.2011.12.003.

Van Koppen, P. J. (2012). Deception detection in police interrogations: Closing in on the context of criminal investigations. *Journal of Applied Research in Memory and Cognition, 1,* 124–125. doi:10.1016/j.jarmac.2012.04.005.

Vredeveldt, A., & Penrod, S. D. (2012). Eye-closure improves memory for a witnessed event under naturalistic conditions. *Psychology, Crime & Law*. doi:10.1080/1068316x.2012.700313.

Vredeveldt, A., & Wagenaar, W. A. (2013). Within-pair consistency in child witnesses: The diagnostic value of telling the same story. *Applied Cognitive Psychology*. doi:10.1002/acp.2921.

Vredeveldt, A., Hitch, G. J., & Baddeley, A. D. (2011). Eyeclosure helps memory by reducing cognitive load and enhancing visualisation. *Memory & Cognition, 39,* 1253–1263. doi:10.3758/s13421-011-0098-8.

Vrij, A. (2006). Challenging interviewees during interviews: The potential effects on lie detection. *Psychology, Crime & Law, 12,* 193–206. doi:10.1080/10683160512331331319.

Vrij, A., & Mann, S. (2001). Who killed my relative? Police officers' ability to detect real-life high-stakes lies. *Psychology, Crime & Law, 7,* 119–132. doi:10.1080/10683160108401791.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition, 1,* 110–117. doi:10.1016/j.jarmac.2012.02.004.

Vrij, A., Akehurst, L., & Morris, P. E. (1997). Individual differences in hand movements during deception. *Journal of Nonverbal Behavior, 21,* 87–102. doi:10.1023/a:1024951902752.

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

Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, R., & Bull, R. (2008b). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior, 32,* 253–265. doi:10.1007/s10979-007-9103-y.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior, 33,* 159–166. doi:10.1007/s10979-008-9143-y.

Vrij, A., Granhag, P. A., & Porter, S. (2010a). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest, 11,* 89–121. doi:10.1177/1529100610390861.

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010b). 'Look into my eyes': Can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law, 16,* 327–348. doi:10.1080/10683160902740633.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20,* 28–32. doi: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,* 579–594. doi:10.1080/1068316x.2010.515987.

Wagenaar, W. A. (2005). De diagnostische waarde van bewijsmiddelen. In M. J. Sjerps & J. A. C. van Voorhout (Eds.), *Het onzekere bewijs: Gebruik van statistiek en kansrekening in het strafrecht* (pp. 3–26). Deventer: Kluwer.

Wagenaar, W. A., & Dalderop, A. (1994). *Remembering the zoo: A comparison of true and false stories told by pairs of witnesses*. Unpublished manuscript, Department of Experimental Psychology, Leiden University.

Wagenaar, W. A., & Van Der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law, 2,* 321–332. doi:10.1080/10683169608409787.

Wagstaff, G. F., Wheatcroft, J. M., Caddick, A. M., Kirby, L. J., & Lamont, E. (2011). Enhancing witness memory with techniques derived from hypnotic investigative interviewing: Focused meditation, eye-closure, and context reinstatement. *International Journal of Clinical and Experimental Hypnosis, 59,* 146–164. doi:10.1080/00207144.2011.546180.

Walczyk, J. J., Mahoney, K., Doverspike, D., & Griffith-Ross, D. (2009). Cognitive lie detection: Response time and consistency of answers as cues to deception. *Journal of Business and Psychology, 24,* 33–49. doi:10.1007/s10869-009-9090-8.

Walczyk, J. J., Griffith, D. A., Yates, R., Visconte, S. R., Simoneaux, B., & Harris, L. L. (2012). LIE detection by inducing cognitive load: Eye movements and other cues to the false answers of "witnesses" to crimes. *Criminal Justice and Behavior, 39,* 887–909. doi:10.1177/0093854812437014.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press.