
Evaluating Truthfulness

► Evaluating Truthfulness: Interviewing and Credibility Assessment

Evaluating Truthfulness: Interviewing and Credibility Assessment

Barry S. Cooper¹, Hugues F. Hervé² and John C. Yuille³

¹The Forensic Alliance, The University of British Columbia, Simon Fraser University and The Forensic Psychiatric Services Commission, Vancouver, BC, Canada

²The Forensic Alliance, Port Moody, BC, Canada

³The Forensic Alliance, The University of British Columbia, Salt Spring Island, BC, Canada

Synonyms

Credibility assessment; Evaluating truthfulness; Fact-finding interviews; Investigative interviewing; Lie/truth detection; Reading verbal and nonverbal behavior; Victim, witness, suspect interviewing

Overview

In the criminal justice context, the overarching goal of the investigative interviewer is to build a case. When facing a truth teller, the interviewer aims to gather enough information to validate that truth. When facing a liar, the interviewer's job is to allow the lie to surface in enough detail to be able to prove its falsehood. In each case, the task relies on a combination of effective interviewing, reading people, and credibility assessment. Are professionals adept at these tasks? If not, why? What does the science suggest that individuals should do to accurately evaluate truthfulness or, at least, improve their ability to evaluate truthfulness? In this entry,

evidence-based and practical answers to these and other questions are offered. In the following sections, terms are defined and methodological problems affecting this complex area are explored, followed by a review of different approaches to evaluating truthfulness. The importance of proper interviewing when one is evaluating truthfulness is subsequently discussed. Finally, the results of evidence-based training in the area are reviewed followed by suggestions for future directions.

Evaluating Truthfulness: Definitions

Evaluating truthfulness requires distinguishing truths from lies. A lie represents a deliberate attempt to deceive without prior notification (Ekman 2001; Vrij 2008). Lies should not be confused with memory errors or other types of non-intentional inaccuracies, which can occur for a variety of reasons unrelated to deceit. There are different types of lies (e.g., falsification, concealment/omission, the incorrect inference dodge, telling the truth falsely) and different topics of deception (e.g., emotional, opinion, factual, intent). In contrast, a truth is an expression of an honestly held belief or the description of a memory that one believes to be the truth, irrespective of its historical accuracy (e.g., narrative vs. historical truth). Thus, in differentiating truths from lies, understanding the intent of the individual is key; while a liar intends to deceive or manipulate another person, the truth teller does not. The result is that the former may experience unique emotional and/or cognitive reactions that the latter does not experience (Cooper et al. 2009). It is these reactions that scientists are trying to understand and practitioners are attempting to identify.

The Difficulty of Evaluating Truthfulness

Most people lie, often a few times per day, and some people lie more than others. Although most lies are trivial in nature (e.g., "great haircut!"),

some have significant personal, social, and legal consequences. Despite the prevalence of deception and the consequences of poor lie detection, most people are not adept at distinguishing truths from lies (Vrij 2008). Indeed, the research, which has typically relied on artificial laboratory paradigms (see below), has generally found that most individuals perform around the level of chance (Bond and DePaulo 2006), irrespective of the nature of their profession or their years of experience in their profession (Ekman and O'Sullivan 1991).

Methodological Problems

Although there is a large literature on evaluating truthfulness, it is marred by a number of problems that impact its generalizability to real-world settings. A major problem with this research is that it has been conducted predominately in the controlled setting of the laboratory. It is argued that, by relying almost exclusively on the laboratory, researchers have committed the offence of methodology (see Yuille 2013). Researchers' strong belief in the utility of controlled research has led them to rely on laboratory analogues to study truthfulness and deceit. More weight has been placed on methodological concerns than on issues concerning generalizability and applicability. In the modal experiment on deception, undergraduate research participants tell the truth or lie about some activity or opinion. The motives to fool others are usually weak (e.g., a small monetary incentive or course credit) and the consequences of being caught in such low-stakes lies have no significant personal or social consequences. The end result is that more is known about how to trigger effects using laboratory designs in undergraduate students instructed to lie under low consequence paradigms than how real-world deception and its detection takes place.

Another methodological problem is the overreliance on group designs. While group designs that compare truth tellers and liars meet stringent research requirements, the practice of evaluating truthfulness focuses on one individual, typically in the context of an interview, and

therefore necessitates a single-subject design for analysis. Furthermore, the field has also been hampered by an overreliance on simplistic statistical procedures, particularly when comparing group averages (i.e., groups of liars are compared to groups of truth tellers) and a small number of indices (i.e., one or two behaviors that may differentiate truth tellers from liars). Using group averages washes out important individual differences in how people reveal their truths and lies – differences that are arguably more important in the real-world practice of evaluating truthfulness (Cooper et al. 2009). For example, if two individuals display their deception in contrasting ways (e.g., one increases and the other decreases a particular behavioral pattern while lying), the average finding would be that the behavior under investigation is unrelated to deception detection. Similarly, looking at only one or two behaviors may result in null findings simply because there are no universal signs of lying and truth telling (see below). That is, the particular behavior(s) may distinguish truths from lies in some participants and not in others. The reduced diagnostic specificity of the behavior(s) may then result in misleading null findings. Consequently, the science and practice of evaluating truthfulness would benefit greatly from more sophisticated methodological and statistical approaches that focus on within-subject methodologies and multivariate analyses.

It is not that laboratory research on evaluating truthfulness is unimportant; indeed, it is important to study how "most people" behave in contexts in which variables can be controlled and manipulated. However, it is also essential to know how to evaluate truthfulness in a particular person in a specified context when the stakes of lying are high (e.g., when the issues are of personal relevance and the consequences are of significance). For this reason, recent efforts have been made to conduct ecologically valid studies, including those examining high-stakes lies (Mann et al. 2002). Clearly, both laboratory and field research on evaluating truthfulness must be conducted to provide converging evidence of the factors revealing truthfulness and deception.

Why Lies Succeed

The current methodological problems of the field notwithstanding, research and practice suggest that lies succeed for a variety of reasons. For example, most lies are difficult to detect as they are often embedded in considerable truth, and some liars are simply too skilled to be detected. Errors in evaluating truthfulness are also made due to faulty assumptions held by the targets of lies (e.g., investigative interviewers), many of which are supported by common myths or through non-evidence-based training. The following is a description of some of the known errors that lead to poor deception detection (for more detail about these and other pitfalls, see Vrij et al. 2010). The most consistent error results from the widely held belief that there is a universal sign or signal diagnostic of deception (i.e., sometimes labelled the *Pinocchio error*; Ekman 2001). Decades of research in the area of evaluating truthfulness, however, have made it clear that there is no universal sign of lying that is displayed by all individuals in all contexts (DePaulo et al. 2003). For example, there has been no empirical support for the proposition offered by proponents of Neuro-Linguistic Programming (NLP) that eye movements are valid indicators of deception (e.g., Wiseman et al. 2012). The misattribution of the emotional state of others – the *Othello error* – is another source of error when evaluating truthfulness (Ekman 2001; e.g., misinterpreting fear of being disbelieved as detection apprehension). When an emotion is displayed, it only signals that the emotion is felt, not the reason for it. Only through effective interviewing can the cause of the emotion be uncovered (Yarbrough et al. 2013). Another error is when the recipient of the lie has little motivation to catch the liar (i.e., sometimes labelled the *ostrich effect*; Vrij 2008). Collusion would be an example of this effect (e.g., when the target wants to believe the liar).

Another error is the *Lie to Me error*, coined after the popular Fox television show, *Lie to Me*, wherein the lead character, played by Tim Roth, is a professed expert lie detector. There are several problems with the approach to evaluating truthfulness depicted in this show. First, Tim

Roth's character presents with an enormous ego associated with overconfidence in his abilities – both of which are known to result in errors in differentiating truths from lies. Second, although he keenly and accurately detects changes in others' behavior – and while detecting change is indeed important to evaluating truthfulness (see below) – he comes to an automatic conclusion of the reason for the behavioral change without considering alternative hypotheses for the observed change (e.g., which also relates to the *Othello error*) or individual differences in behavior (note: the latter has also been called the *idiosyncrasy error*; Ekman 2001). A failure to consider multiple hypotheses for what is observed and heard coupled with impulsive decision making increases the chance of errors. Third, the fictional character believes that mere passive observation is a valid avenue to evaluate truthfulness accurately, shunning the notion that interviewing is important. As discussed below, however, evaluating truthfulness is heavily dependent on effective interviewing. The *Lie to Me error* concerns, in part, a combination of other errors and arguably represents a higher-order error. This error may be the most difficult to extinguish as, at some level, it involves a lack of core critical thinking skills.

Approaches to Evaluating Truthfulness

Currently, a variety of approaches to evaluating truthfulness have gained attention and/or empirical support. These approaches rely either on psychophysiological measures or the use of observational techniques. Each is discussed in turn.

Psychophysiological Techniques

A number of technologies based on psychophysiological assessments have been applied to the area of evaluating truthfulness. The polygraph is one of the most long-standing psychophysiological tools used to evaluate truthfulness in the modern era, and it has been the subject of an extensive amount of research (see Griesel and Yuille 2007). There are a number of different ways in which the polygraph is used, including

the Directed Lie Test, the Control Question Test, the Guilty Knowledge/Concealed Knowledge Test, and the Relevant-Irrelevant Test. The polygraph is not a “lie detector,” although it has been referred to as such, often in the popular media. Rather, the polygraph detects changes in physiological arousal through measurements of heart rate, galvanic skin conductance, respiration, etc. Such changes may be due to lying but may also be due to other factors (e.g., stress, fear of being disbelieved, increased cognitive load). There are a number of concerns about the polygraph’s false-positive and false-negative rates, which, in most countries, precludes its use in contexts such as employee screening (National Research Council 2003). Similarly, concerns about the validity and reliability of the polygraph have also led to restrictions on its use and applicability in relation to matters of the criminal justice system. Arguably the strongest facet of the polygraph is the administrator. Anecdotally, effective polygraphers are excellent interviewers and critical thinkers – they treat the physiological results of the polygraph as simply one piece of data in the larger puzzle.

Recently, some companies have developed voice stress analyzers, which detect changes in verbal characteristics (e.g., pitch and tone of voice). Although these companies have marketed these tools as highly effective lie detectors, the research has shown these marketing claims to be false (Dampousse 2008). Although detecting change in behavior (e.g., via vocal clues) is essential to evaluating truthfulness (see below), such changes may be due to lying or to a host of other reasons (Cooper et al. 2009). The reason for the change can only be uncovered via testing alternative hypotheses through effective interviewing.

There has also been recent interest in the use of thermal imaging technologies, which detect body heat changes, in the evaluation of truthfulness. To date, there has been no strong support for thermal imaging as a valid method of evaluating truthfulness (Warmelink et al. 2011); however, the research has been sparse. Electroencephalography (EEG), a measure of neural activity, typically via event-related potentials (ERPs), has also been used in the study of evaluating truthfulness

and also with limited success (Rosenfeld 2002). The past decade or so has seen greater efforts to identify the neural processes involved in deception via functional magnetic resonances imaging (fMRI) technology. Although gains have been made in this field, there is no empirical foundation for the notion that fMRI can be used to reliably evaluate truthfulness (Spence 2008). Indeed, although certain areas of the brain (i.e., within the frontal and parietal lobes) may be more active than others when deception occurs, research indicates that fMRI technology does not reveal neural processes in these areas that are unique to deceit (Monteleone et al. 2009).

There is a pattern that emerges from an examination of the various psychophysiological techniques proposed to evaluate truthfulness – as far as is currently known, there is no psychophysiological response that is unique to deception in all individuals and in all contexts. A failure to understand this fact leads individuals to commit the *Pinocchio error* (see above). Figuratively and, in some cases, literally, most of the psychophysiological measures detect *hot spots* (see below) – a change from baseline and/or inconsistencies across channels measured (Cooper et al. 2009) – not lies. Although detecting hot spots is important, such change is not diagnostic of lying. Evaluating truthfulness is a two-step process: first the observer needs to detect behavioral changes and/or inconsistencies; second, the observer needs to determine the reason the behavior occurred.

Observational Techniques: Reading Verbal and Nonverbal Behavior

Observational techniques for evaluating truthfulness rely on reading verbal and nonverbal behavior associated with truth telling and lying. Relative to psychophysiological technologies, these techniques – not dependent on any equipment – are low cost, transportable, generalizable, and noninvasive. There are two clear “camps” of researchers in the observational study of evaluating truthfulness: those in the cognitive camp (i.e., primarily examining verbal behavior) and those in the emotional camp (i.e., primarily examining nonverbal behavior). Most researchers in the

cognitive camp ascribe to the theoretical assumption that, all factors considered equal, lying is more cognitively demanding than telling the truth because the liar has to fabricate details that do not exist, remember his/her lie, suppress information, evaluate the success of the lie, adapt accordingly, etc., while the truth teller simply has to recall and report details from memory (DePaulo et al. 2003; Vrij et al. 2010). Researchers from this camp believe that the increase in cognitive load (i.e., mental effort) associated with deception leads to behavioral changes that betray the liar, such as changes or particularities in verbal content, verbal style, and body language.

Most researchers in the emotional camp believe that lying may result in an emotional response (e.g., fear – detection apprehension; happiness – duping delight) and, as a consequence, emotional “leakage” – that is, the display of an emotion that may leak out and thus betray a lie (Ekman 2003). Micro and subtle behavior are viewed as especially important, with the former being a full behavioral display (e.g., a full two-sided shrug or a complete display of anger in the face) that occurs for only a fraction of a second and the latter being a partial behavioral display (e.g., a shrug on only one side; anger displayed only in the eyes) that can last a fraction of a second or longer. These types of behaviors are involuntary and therefore provide messages that the individual is trying to hide, be it to themselves or the interviewer. This is in contrast to macro behaviors, which are long-lasting full displays of behaviors that are more under control (e.g., can be faked by a deceptive person). To date, most of the research has focused on displays of emotional leakage in the face, particularly micro and subtle expressions, and changes in body language and vocal characteristics. This research focuses on identifying both changes in behavior, as change does not occur randomly, and inconsistencies between the various behavioral channels, such as claiming to love someone but showing disgust or contempt when talking about them.

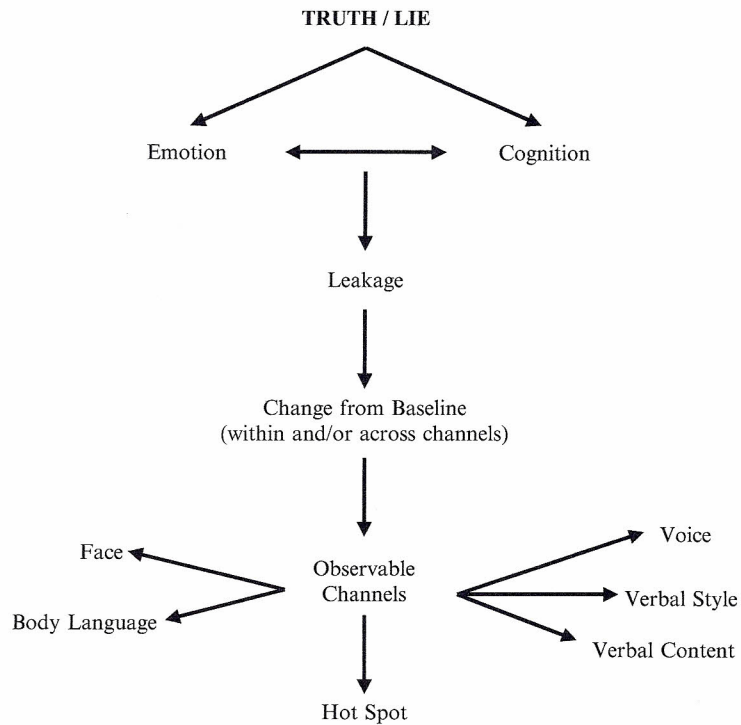
In recent years, researchers from both camps have started to believe that their respective

viewpoints and lines of research are complementary rather than contradictory. Most researchers and practitioners now agree that telling the truth or a lie may result in emotional and/or cognitive leakage, which itself may lead to changes in verbal and/or nonverbal behavior, and that the ability to trigger, identify, and make sense of these behavioral changes relies on effective interviewing strategies (Vrij et al. 2010). An integrated stepwise method to evaluating truthfulness which incorporates the evidence-based views of both camps and stresses the importance of effective interviewing during the collection of information has recently been provided (i.e., Cooper et al. 2009). First, according to this method, the interviewer must assess the interviewee’s baseline behavior (i.e., the individual’s “normal” way of behaving when not lying; e.g., via prior recordings, via the rapport stage of an investigative interview). Second, the interviewer must actively listen and observe for changes from baseline within a variety of observable behavioral channels (i.e., face, body, voice, verbal style, and verbal content; see Fig. 1), as well as inconsistencies across these channels. As depicted in Fig. 1, these changes and/or inconsistencies reflect cognitive and/or emotional load that has leaked out in behavior. This leakage should not be impulsively interpreted as a sign of deception given that there is no *Pinocchio* response. Rather, observed leakage should be viewed as a hot spot, that is, a behavior of importance that may be due to a number of causes (e.g., related to emotional or cognitive leakage) of which lying is only one possibility. In other words, the interviewer should consider multiple hypotheses for what she/he sees or hears. As shown in Fig. 1, the interviewer should also focus equally on signs of truth telling and lying, as focusing on either in isolation predisposes one to make the error of believing the liar or disbelieving the truth teller, respectively.

Finally, once a hot spot has been identified, the interviewer must test his/her hypotheses for the hot spot. This is accomplished by further exploring the topic that triggered the hot spot via effective interviewing (see below) and/or by employing specific interviewing techniques designed to distinguish truth tellers from liars.

Evaluating Truthfulness: Interviewing and Credibility Assessment,

Fig. 1 Model for
evaluating truthfulness
(From Cooper et al. 2009)



For example, the interviewer, after noting a hot spot, should allow the interviewee to continue with his/her statement. Once the statement is complete, the interviewer can return to the topic that led to the initial hot spot. If another hot spot arises, it provides more evidence that the hot spot was related to the topic as opposed to another reason. This final step is crucial as it permits the interview to rule out hot spots unrelated to lying, thereby allowing the truth teller to demonstrate his/her credibility; or it could result in further hot spots (and therefore evidence) in the liar that betrays his/her deceptiveness. An essential feature of this approach is that the interviewer must delay judgement until enough information has been collected and alternative hypotheses have been ruled out through effective interviewing. Prior to outlining the basics of effective interviewing, each behavioral channel noted above is briefly reviewed.

Face

Although a number of behavioral channels can display emotional leakage, the face is the clearest

channel for this type of leakage; it is also the most researched behavioral channel concerning emotional leakage (Ekman 2001; Porter et al. 2012). As noted above, although identifying macro facial expressions can be informative (e.g., happiness, sadness, fear, surprise, anger, contempt), the detection of micro and subtle facial expressions is most revealing as these often signal information that the interviewee is trying to hide. This is especially true when they contradict what is said. For example, if a husband suspected of a domestic homicide claims to have had a loving relationship with his wife but concurrently shows a subtle expression of anger, this is an important hot spot that needs to be further investigated.

In addition to signalling emotional reactions, the face is a potential source of information about cognition (e.g., the pursing of the lips typically signals a mental search for information). For example, certain facial movements (e.g., the lowering of the eyebrows) are signs of concentration and, therefore, reflect increased cognitive load. If an interviewee shows these signs, the reason for the behavioral change may be obvious

(e.g., responding to a relatively difficult question, attempting to recall a distant memory) or it may be a clue that an interviewee may be thinking too hard relative to the level of difficulty of the question (e.g., when asked his/her name or birthday), thereby suggesting that one hypothesis to consider is that the individual is being deceptive. Interviewers must entertain multiple hypotheses for the observed behavioral change in the face (e.g., the lowering of the eyebrows could be related to concentration or subtle anger; Ekman 2003). If the observed behavioral change is out of context, it is a hot spot (Cooper et al. 2009).

Body Language

Body movement or change is also an important behavioral channel to read when evaluating truthfulness. Indeed, certain activities (e.g., sweating, fidgeting, finger tapping) may provide clues to one's emotional state (e.g., stress, nervousness). There are also gestures that individuals display that are meaningful clues to other emotions or cognitions, particularly when viewed as a change from baseline. For example, the literature has identified three different types of gestures that are important to consider when reading people and evaluating truthfulness: emblems, manipulators, and illustrators (e.g., Friesen et al. 1979). Emblems are literally a type of body language (e.g., nodding one's head "yes") and are completely culturally determined. For example, the "thumbs up" sign has an entirely different meaning in North America than it does in some contexts in the Middle East (i.e., it is viewed as a rude gesture). Knowledge of the specific emblems of a culture is required to understand their meaning. The display of emblems is usually voluntary but emblematic slips may involuntarily leak information (e.g., when someone shrugs their shoulders – a sign of uncertainty – while attempting to convey a verbal message of certainty) – emblematic slips are thus, by definition, hot spots.

Manipulators, also referred to as self-manipulations, refer to the movement of one body part on another (e.g., scratching one's nose or ear, playing with one's hair; there are self-manipulators and object manipulators). Some research has shown that deception is associated

with an increased use of manipulators (Porter et al. 2008). When evaluating truthfulness in practice, it is important to pay attention to individuals' baseline display of manipulators as a change from baseline is considered a hot spot (Cooper et al. 2009). Some individuals increase their use of manipulators when under increased emotional arousal, yet others may show an increase in manipulators when relaxed (e.g., an increase in grooming behavior).

Illustrators are hand movements used to illustrate speech as it is being spoken. Although some research (see DePaulo et al. 2003) suggests that individuals tend to decrease their use illustrators when being deceptive (e.g., as such may signify an increase in cognitive load), it is important to consider a person's baseline. Some people show an increase in illustrators when their cognitive load increases; others show a decrease. A change from baseline (i.e., an increase or decrease) would signify a hot spot. It is the interviewer's task to figure out the reason for the change.

Voice

The voice is another behavioral channel in which leakage can be observed as assessed via pitch, tone, loudness, etc. For example, a vocal change could be a cue to individuals' emotional state with the voice tending to have an edge when anger is experienced and to be higher in pitch when fear or anger is felt (Ekman 2003). It has also been shown that peoples' voices tend to soften when being deceptive; however, it has also been demonstrated that the voice tends to soften when people are uncertain or feeling sad. Thus, as with the other behavioral channels, a multiple hypothesis testing approach must be used to delineate the reason for the vocal change. The change is meaningful but should not be interpreted as a lie per se, as the companies that market voice stress analyzers suggest. For example, if an interviewee softens his voice relative to his baseline while being interviewed during an assessment of his suicide risk, the softening could suggest sadness (e.g., a potential risk factor for suicide), deception, or some other factor. Only through effective interviewing could the reason for the expressed vocal change be uncovered.

Verbal Style

Verbal style is an aspect of language that can leak both emotional and cognitive load. For example, relative to an individual's baseline verbal style, an increase in pauses or filled pauses could suggest an increase in cognitive load (Mann et al. 2002). The same would be the case for word repetitions, as well as an increase in stuttering or the use of jargon. It is the change in verbal style – relative to baseline – that is important, as opposed to the verbal style per se. For example, if an interviewee's baseline verbal style is the use of jargon (e.g., “you know what I mean,” “you know”) and the interviewee uses such jargon throughout an interview, the displayed jargon is relatively meaningless in terms of evaluating truthfulness. If, however, an interviewee does not typically use such jargon but does in relation to answering critical questions during an investigative interview, the change in verbal style should be considered a hot spot. Pronoun usage is another aspect of verbal style that has been applied to the area of evaluating truthfulness. For example, it has been empirically demonstrated that, all other factors considered equal, individuals tend to use less first person pronouns (e.g., I, my, me) when lying than when telling the truth, arguably in an attempt to distance themselves from their lies. As with the other behavioral channels, it is the change in pronoun usage – relative to baseline – that is important, not the rate of pronoun usage per se.

Verbal Content

A considerable amount of research has examined verbal content as a behavioral channel in which to evaluate truthfulness. A number of techniques have been developed to analyze the quality and quantity of verbal content including Scientific Content Analysis (SCAN; Sapir 1987, 2000), Reality Monitoring (RM; Johnson and Raye 1981), and Criteria-Based Content Analysis (CBCA; Yuille 1988). SCAN is a widely used technique developed by a former Israeli polygrapher. It involves eliciting a written statement about an event (e.g., an alleged crime) followed by an analysis of the statement with the use of

certain criteria (e.g., out of sequence, social introductions). Anecdotal evidence of its utility notwithstanding, there has been no empirical support for SCAN (Nahari et al. 2012). Moreover, certain assumptions underlying SCAN are theoretically inconsistent with known memory/psychological processes.

RM is based on the theoretical assumption that memories of experienced events will contain more contextual information as well as more external-sensory information than memories that are the product of fabrication/imagination (Johnson and Raye 1981). The RM technique involves an interviewee rating their own memory (e.g., for vividness, detail) followed by an external rater who follows the same procedure after reading the interviewee's transcript. While RM is more firmly grounded in science than SCAN and there has been some evidence attesting to its utility, it is not without its problems, particularly when applied to the forensic arena (Colwell et al. 2013).

CBCA is one of the most researched techniques applied to the evaluation of the credibility of verbal content (Vrij 2005). It is one part of a larger system – Statement Validity Analysis (Yuille 1988). Of note, it is not possible to apply CBCA to an interview (or a transcript of an interview) unless the interview is of high quality (e.g., a free narrative is required for analysis). In some respects, the theoretical assumptions underlying CBCA are similar to those of RM. CBCA is based on the Undeutsch hypothesis that posits that memories of actual experiences differ in quantity and quality from memories based on invention (Undeutsch 1989). A number of versions of CBCA are in existence but most contain at least 19 criteria that are applied to an interview /statement/transcript. Depending on the version, the criteria are divided into the following five areas which are applied to a statement by a trained rater: general characteristics (i.e., coherence, spontaneous reproduction, appropriate amount of detail), specific content (e.g., contextual embedding, descriptions of interactions), peculiarities of the content (e.g., unusual details, related external associations), motivational

features (e.g., spontaneous corrections, self-deprecation), and stylistic features (e.g., theme-related changes, rigid repetition). The use of CBCA has been empirically shown to be able to differentiate deceptive from truthful statements (i.e., concerning memories of past events) at a rate ranging from 55 % to 90 % in samples of children, adult witnesses, and adult offenders.

There are a number of misperceptions about CBCA that have implications for its applied utility. For example, CBCA is a complex qualitative procedure, not a quantitative tool (see Griesel et al. 2013). Indeed, it is not a matter of simply adding up the criteria present in a statement to make a judgement about its credibility as some research would suggest. Rather, as certain criteria arguably should have more weight than others (e.g., appropriate amount of detail vs. unusual details), the trained rater should apply CBCA in a qualitative manner. Further, training matters – in order to apply the procedure effectively, a considerable amount of training is required. This would include acquiring knowledge about the factors that impact memory (see Hervé et al. 2013) in order to know what is considered an appropriate amount of detail in a statement.

Summary of Observational Techniques

Each of the five behavioral channels can provide useful information for evaluating truthfulness. The channels must be considered together and not in isolation. Often, the most telling hot spots occur when there is an inconsistency between behavioral channels (e.g., when an interviewee says “no” but nods his head “yes”). These and other hot spots allow the interviewer to navigate the interview to identify areas in need of further enquiry. The interviewer’s job is to collect sufficient information to validate the credibility of the truthful person or elicit sufficient information to reveal the lie in the deceptive individual. Only when an interview provides a high quality and quantity of details can such tools as CBCA be applied. As no tool or technique is foolproof, it is imperative that the investigative interviewer corroborate his/her conclusion (e.g., verify alibis/whereabouts).

The Importance of Effective Interviewing When Evaluating Truthfulness

The importance of an effective interview when evaluating truthfulness cannot be overstated. In practice, lies are embedded in a great deal of truths and the ability to differentiate what is truthful from fictional relies heavily on effective interviewing. A poor interview only serves to further blur the distinction between truth and deception. The overarching goal of an effective interview is to gather uncontaminated information, that is, the interviewee’s version of events. In other words, the interviewer’s task is to cue memory, not lead memory. Leading questions run the risk of causing the interviewee to unintentionally provide inaccurate information, which could be misinterpreted as a sign of deception. Leading or suggestive questions can also unintentionally telegraph the interviewer’s goal or suspicions, thereby allowing deceivers to alter their deceptive tactic(s). In contrast, non-leading, open-ended questions allow the interviewee to provide what she/he knows in a spontaneous manner. Spontaneity is one of the most potent signs of truthfulness (Colwell et al. 2013). In order to accomplish the goal of eliciting uncontaminated information, it is suggested that interviewers adopt a memory-based approach to interviewing (Yarbrough et al. 2013), such as the Step-Wise Interview Guidelines (Yuille et al. 2009), the Cognitive Interview (Fisher and Geiselman 1992), or the PEACE model of interviewing (Milne and Bull 1999), all of which have the goal of cuing and exhausting memory without contamination.

A memory-based approach is crucial given that the typical investigative interview focuses on an interviewee’s memory about some past event. Accordingly, interviewers should be knowledgeable about how memory works, what factors impact memory, etc., in order to effectively navigate an interview and accurately evaluate the truthfulness of the memory received (see Hervé et al. 2013 for a review of memory patterns and of the biopsychosocial predisposing,

precipitating, and perpetuating factors influencing memory). That is, the more one knows about memory, the easier it is to elicit it and evaluate its credibility. Indeed, the quality and quantity of details contained in a reported memory are key determinants to its credibility, and the “appropriateness” of these details in a particular case depends on the interviewees’ predisposing, precipitating, and perpetuating memory factors.

While there are different memory-based interview models, they all share the following features. The first feature of an effective interview is preparation. This involves, among other factors, gathering as much background information about the interviewee and the precipitating event (s) as possible. The more the interviewer knows about the interviewee and the subject under investigation, the better she/he is equipped to develop rapport, assess baseline, identify and assess hot spots, entertain alternative hypothesis, and exhaust memory. For example, preparation may reveal culturally specific information that may help in the interpretation of body language, while knowledge of relevant crime patterns and the evidence in the case at hand may reveal inconsistencies between this information and the interviewee’s statement. The second feature is rapport building. During the initial stages of an interview, the interviewer should develop a working relationship with the interviewee that assists in placing the interviewee at ease and encourages dialogue. During this part of the interview, the interviewer should assess the interviewee’s baseline (a process that continues throughout the interview). The better the rapport, the clearer the baseline and the easier it is to detect deviations from baseline (or hot spots). For example, if the interviewer projects suspiciousness or dislike of the interviewee, this could cause stress or anxiety in the interviewee. This emotional load will leak out in behavior which could either be misinterpreted as a sign of lying or serve to mask leakage associated with deception to critical questions (i.e., the signal becomes lost in the noise). This is why it is recommended that interviewers adopt, at least initially, an information-gathering style of asking questions (i.e., as opposed to an accusatory style

of interviewing – challenges are best left to the later stages of interviews).

The third feature of an effective interview is that it allows the interviewee every opportunity to provide his/her version of events. As noted above, this involves cuing the interviewee’s memory via non-leading, open-ended questions. This typically involves a funnel approach to interviewing wherein the interviewer attempts to elicit an uninterrupted free narrative followed by open-ended W-H questions before asking specific questions as deemed necessary. Interruptions are avoided as they disrupt the reconstructive nature of memory. This also often serves to reduce the amount of questions being asked; the fewer the questions needed to exhaust memory, the better the quality of the information provided. Indeed, research has demonstrated that interviewees provide more correct information about past events during the free narrative aspect of interviewing than in response to specific questions. Further, a free narrative is the ideal aspect of an interview for the application of certain verbal credibility assessment techniques such as CBCA.

In addition to sharing the aforementioned features, memory-based interviewing approaches also include various strategies (e.g., narrative repetition, different perspective, backward recall) for enhancing memory when the provided memory appears to lack an appropriate amount of detail. There are many potential reasons why someone may provide a poorly detailed memory (see Hervé et al. 2013), some of which are valid (e.g., time-related memory decay) while others may be deceptive (e.g., making up an event to conceal the truth). Memory-enhancing techniques rely on the reconstructive and cued nature of memory, which dictate that one’s narrative should appear somewhat different from one recall to the next (e.g., a few details added and/or a few details omitted), that additional details can be elicited given the right cue, and that the manner in which the information is recalled can be used to cue additional details (e.g., field vs. observer perspective; forward vs. backward recall). These memory-enhancing strategies are therefore important tools for an interviewer in his/her

attempt to clarify the possible reason(s) for the observed lack of detail (i.e., test alternative hypotheses). On the one hand, the truthful person – having a real memory to rely on – should provide additional information when asked, for example, to repeat his/her narrative (irrespective of the perspective or direction of recall; note: this information could serve to validate his/her memory and/or provide additional case-relevant details). On the other hand, the liar who fabricated a story and therefore has no memory to rely on (unless the lie is based on a true experience) is predisposed to provide either a rigid repetition and/or a poverty of additional details, both of which are notable hot spots. Moreover, some of these techniques require additional cognitive resources, which liars find difficult to manage. For example, all other factors considered equal, it is easier for an interviewee to recall an event in chronological order in comparison to backward order. Research has found that the increase in cognitive load is more pronounced for liars than truth tellers and that the resulting leakage facilitates the task of differentiating truths from lies (Vrij et al. 2008). Recent research efforts have focused on defining additional interview strategies that create greater cognitive load in truth tellers than in liars, including asking unanticipated questions and the strategic disclosure of evidence by interviewers (Jordan et al. 2012; Vrij et al. 2009). Regardless of what interviewing strategies are used, the resulting leakage should be viewed as a hot spot. Multiple hypotheses must be entertained for the reason for the elicited hot spot, of which deception is only one possibility.

Evidence-Based Training on Evaluating Truthfulness

As indicated above, without any training, it has been shown that the average person evaluates truthfulness at the level of chance (Bond and DePaulo 2006; Ekman and O'Sullivan 1991; Vrij 2008). A number of researchers/organizations have provided evidence-based training (i.e., based on the principles of reading

verbal and nonverbal behavior discussed above) to various professional groups (e.g., law enforcement, corrections) in the general area of detecting deceit/evaluating truthfulness. However, there exist only a few studies examining the effectiveness of such training. In general, these studies show that 1- to 2-day workshops can improve participants from about chance level at pretest to an accuracy rate of 77 % or higher at posttest (e.g., Porter et al. 2000; Colwell et al. 2009). Training gains, although smaller in magnitude, have also been demonstrated after 3-hour workshops (e.g., Colwell et al. 2012: 58 % at pretest to 72 % at posttest; Porter et al. 2010: 51 % at pretest to 61 % at posttest). While these findings are encouraging, it remains to be seen whether the gains made in these workshop generalize to the real world. Clearly, future research on the effectiveness and generalizability of evidence-based training is required.

The gains that individuals make in training notwithstanding, it is important to acknowledge that the field will never advance to the point that training will provide the knowledge and skills to allow judgements to be 100 % accurate. Evaluating truthfulness is far too complex of an enterprise. It is more realistic to be able to improve most individuals' abilities to evaluate truthfulness from chance level to some reliable level above chance. The goal should not be to produce *wizards* – rare individuals who appear to be able to reliably evaluate truthfulness at an exceptionally high rate (O'Sullivan 2013). That said, the continued study of these types of individuals will hopefully lead to new insights into the personality characteristics of effective evaluators, the process of differentiating truths from lies, and key features to consider.

Summary and Future Directions

This entry provided an overview of the nature and complexity of evaluating truthfulness. Common pitfalls and approaches were described, and the interaction between effective interviewing, reading verbal and nonverbal behavior, and assessing credibility was reviewed. It was shown that

individuals can learn to improve their ability to evaluate truthfulness through evidenced-based training. Despite the gains made in the field through decades of social science research, more research is clearly needed, particularly research that uses single-subject designs in real-world settings and takes into consideration the perspectives of practitioners who evaluate truthfulness as a core aspect of their professional duties. With a single-subject design, an individual's baseline could be assessed and changes from baseline could be examined within and across all of the observable behavioral channels (i.e., face, body language, voice, verbal style, and verbal content). In this regard, more attention should be devoted to individual differences (e.g., personality, personality pathology, mental illness, cognitive functioning, emotional intelligence, culture) as these differences can affect each aspect of the process of evaluating truthfulness reviewed above.

It is recommended that high-stakes lies be examined in a variety of contexts. For example, a promising area of research in evaluating truthfulness concerns the study of computer (and related)-mediated communication (CMC; Hancock and Woodworth 2013). Future research should also examine the ability of individuals to evaluate truthfulness in the context of interviewing. Variables such as the length of the interview, the focus of the interview (e.g., memory for past vs. future actions), and the nature of the interview (e.g., overt vs. covert) should be examined in the laboratory and in the field (e.g., low- vs. high-stakes lies). As well, research should identify the type of individuals best geared to evaluate truthfulness – no doubt the research on wizards holds promise in this regard. Not everyone is suited to be an investigative interviewer and/or able to effectively evaluate truthfulness. Key characteristics are likely to include an information-gathering mind-set, critical thinking skills, and the capacity to develop rapport, read people, and make decisions about truthfulness holistically.

Finally, research attention should be devoted to deciphering the most effective way to deliver evidence-based training to ensure that the

training results generalize to on-the-job performance. Training should focus on investigative interviewing as a foundation (e.g., how to gather information – first level training), followed by modules on reading people (e.g., identifying and probing for hot spots – second level training), and evaluating truthfulness (e.g., making sense of hot spots via hypothesis testing and verbal content analysis – third level training). Actual cases/recorded interviews should be used for demonstration and practice as part of the training. Ideally, the training would include a refresher component – which has shown promise in the interviewing training literature (Price and Roberts 2011). This could include literature updates and practice cases to avoid drift in training gains. Finally, post-training on-the-job performance should be assessed in the context of supervision/mentoring and the review of actual recorded interviews in which ground truth has been determined.

Related Entries

- ▶ Behavioral Investigative Advice
- ▶ British Police
- ▶ Criminal Investigative Analysis
- ▶ Detecting Deception with fMRI
- ▶ Evidence-Based Policing
- ▶ False Confessions and Police Interrogation
- ▶ False Memories
- ▶ Interview and Interrogation Methods Effects on Confession Accuracy
- ▶ Interviewing Eyewitnesses

Recommended Reading and References

- Bond CF, DePaulo BM (2006) Accuracy of deception judgements. *Personal Soc Psychol Rev* 10:214–234
- Colwell K, Hiscock-Anisman C, Memon A, Colwell LH, Taylor L, Woods D (2009) Training in assessment criteria indicative of deception (ACID) to improve credibility assessment. *J Forensic Psychol Pract* 9:199–207
- Colwell LH, Colwell K, Hiscock-Anisman C, Hartwig M, Cole L, Werdin K, Youschak K (2012) Teaching professionals to detect deception: the efficacy of a brief training workshop. *J Forensic Psychol Pract* 12:68–80

- Colwell K, Hiscock-Anisman C, Jacquelyn Fede J (2013) Assessment criteria indicative of deception (ACID): an example of the new paradigm of differential recall enhancement. In Cooper BS, Griesel D, Ternes M (eds) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. Springer, New York, pp 259–291
- Cooper BS, Hervé H, Yuille JC (2009) Evaluating truthfulness: detecting truths and lies in forensic contexts. In: Bull R, Valentine T, Williamson T (eds) Handbook of psychology of investigative interviewing. Wiley-Blackwell, Chichester, pp 301–328
- Damphousse KR (2008) Voice stress analysis: only 15 percent of lies about drug use detected in a field test. *NIJ J* 258:8–12
- DePaulo BM, Lindsay JL, Malone BE, Muhlenbruck L, Charlton K, Cooper H (2003) Clues to deception. *Psychol Bull* 129:74–118
- Ekman P (2001) Telling lies: clues to deceit in the marketplace, politics, and marriage. W.W. Norton & Co., New York
- Ekman P (2003) Emotions revealed: recognizing faces and feeling to improve communication and emotional life. Henry Holt and Company, New York
- Ekman P, O'Sullivan M (1991) Who can catch a liar? *Am Psychol* 46:913–920
- Fisher RP, Geiselman RE (1992) Memory-enhancing techniques for investigative interviewing: the cognitive interview. Charles C. Thomas, Springfield
- Friesen WV, Ekman P, Wallbott H (1979) Measuring hand movements. *J Nonverbal Behav* 4(2):97–112
- Griesel D, Ternes M, Schraml D, Cooper BS, Yuille JC (2013) The ABC's of CBCA: verbal credibility assessment in practice. In Cooper BS, Griesel D, Ternes M (eds) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. Springer, New York, pp 293–323
- Griesel D, Yuille JC (2007) Credibility assessment in eyewitness memory. In: Togliani MP, Read JD, Ross DF, Lindsay RCL (eds) Handbook of eyewitness psychology. Volume I, memory for events. Lawrence Erlbaum, Mahwah, pp 339–370
- Hancock JT, Woodworth M (2013) An “Eye” for an “I”: the challenges and opportunities for spotting credibility in a digital world. In Cooper BS, Griesel D, Ternes M (eds) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. Springer, New York, pp 325–340
- Hervé HF, Cooper BS, Yuille JC (2013) Biopsychosocial perspectives on memory variability in eyewitnesses. In Cooper BS, Griesel D, Ternes M (eds) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. Springer, New York, pp 99–142
- Johnson MK, Raye CL (1981) Reality monitoring. *Psychol Bull* 88:67–85
- Jordan S, Hartwig M, Wallace B, Dawson E, Xihani A (2012) Early versus late disclosure of evidence: effects on verbal clues to deception, confessions, and lie catcher' accuracy. *J Investig Psychol Offender Profiling* 9:1–12
- Mann S, Vrij A, Bull R (2002) Suspects, lies and videotape: an analysis of authentic high-stake liars. *Law Hum Behav* 26:365–372
- Milne R, Bull R (1999) Investigative interviewing: psychology and practice. Wiley, Chichester
- Monteleone GT, Phan L, Nusbaum HC, Fitzgerald D, Irick J-S, Fienberg SE, Cacioppo JT (2009) Detection of deception using fMRI: better than chance but well below perfection. *Soc Neurosci* 4(6):528–538
- Nahari G, Vrij A, Fisher R (2012) Does the truth come out in the writing? SCAN as a lie detection tool. *Law Hum Behav* 36(1):68–76
- National Research Council (2003) The polygraph and lie detection. Committee to review the scientific evidence on the polygraph. The National Academic Press, Washington
- O'Sullivan M (2013) Is *le mot juste*? The contextualization of words by expert lie detectors. In Cooper BS, Griesel D, Ternes M (eds) Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. Springer, New York, pp 239–257
- Porter S, Doucette NL, Woodworth M, Earle J, MacNeil B (2008) Half the world knows how the other half lies: investigation of verbal and non-verbal signs of deception exhibited by criminal offenders and non-offenders. *Leg Criminol Psychol* 13:27–38
- Porter S, Juodis M, ten Brinke L, Klein R, Wilson K (2010) Evaluation of a brief deception detection training program. *J Forensic Psychiatry Psychol* 21:66–76
- Porter S, ten Brinke L, Wallace B (2012) Secrets and lies: involuntary leakage in deceptive facial expressions as a function of emotional intensity. *J Nonverbal Behav* 36:23–37
- Porter S, Woodworth M, Birt AR (2000) Truth, lies and videotape: an investigation of the ability of federal parole officers to detect deception. *Law Hum Behav* 24:643–658
- Price HL, Roberts KP (2011) The effects of an intensive training and feedback program on police and social workers' investigative interviews of children. *Can J Behav Sci* 43(3):235–244
- Rosenfeld JP (2002) Event-related potentials in the detection of deception, malingering, and false memories. In: Kleiner M (ed) Handbook of polygraph testing. Academic, San Diego, pp 265–286
- Sapir A (1987, 2000) The LSI course on scientific content analysis (SCAN). Laboratory for scientific interrogation, Phoenix, AZ
- Schacter DL (1996) Searching for memory. Basic Books, New York
- Spence SA (2008) Playing devil's advocate: the case against fMRI lie detection. *Leg Criminol Psychol* 13:11–26
- Undeutsch U (1989) The development of statement reality analysis. In: Yuille JC (ed) Credibility assessment. Kluwer Academic Publishers, Dordrecht, pp 101–120

- Vrij A (2005) Criteria-based content analysis: a qualitative review of the first 37 studies. *Psychol Public Policy Law* 11(1):3–41
- Vrij A (2008) *Detecting lies and deceit: pitfalls and opportunities*, 2nd edn. Wiley, Chichester
- Vrij A, Granhag A, Porter S (2010) Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychol Sci Public Interest* 11(3):89–21
- Vrij A, Leal S, Granhag PA, Mann S, Fisher RP, Hillman J, Sperry K (2009) Outsmarting the liars: the benefit of asking unanticipated questions. *Law Hum Behav* 33:159–166
- Vrij A, Mann S, Fisher R, Leal S, Milne B, Bull R (2008) Increasing cognitive load to facilitate lie detection: the benefit of recalling an event in reverse order. *Law Hum Behav* 32:253–265
- Warmelink L, Vrij A, Mann S, Leal S, Forrester D, Fisher R (2011) Thermal imaging as a lie detection tool at airports. *Law Hum Behav* 35:40–48
- Wiseman R, Watt C, ten Brinke L, Porter S, Couper S-L, Rankin C (2012) The eyes don't have it: lie detection and neuro-linguistic programming. *PLoS ONE* 7(7):1–5
- Yarbrough J, Hervé HF, Harms R (2013) The sins of interviewing: errors made by investigative interviewers and suggestions for redress. In Cooper BS, Griesel D, Ternes M (eds) *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment*. Springer, New York, pp 59–95
- Yuille JC (1988) The systematic assessment of children's testimony. *Can Psychol* 29(3):247–262
- Yuille JC (2013) The challenge for forensic memory research: methodology. In Cooper BS, Griesel D, Ternes M (eds) *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment*. Springer, New York, pp 3–18
- Yuille JC, Cooper BS, Hervé HF (2009) The step-wise guidelines for child interviews: the new generation. In Casonato M, Pfafflin (eds) *Pedoparafiile: psychological perspectives, forensic psychiatric*, Italy: Franco Angeli

Evaluation

- ▶ Bullying Prevention: Assessing Existing Meta-Evaluations

Event Characteristics

- ▶ Estimator Variables and Eyewitness Identification

Event Control: Criminal Event, Routines Activities, Presentness, Occasions

- ▶ Control Theory

Evidence-Based Corrections

- ▶ Examining the Effectiveness of Correctional Interventions

Evidence-Based Crime Policy

- ▶ Evidence-Based Policy in Crime and Justice

Evidence-Based Policing

Cynthia Lum and Christopher S. Koper
Department of Criminology, Law and Society,
George Mason University, Fairfax, VA, USA

Overview

Evidence-based policing is a law enforcement perspective and philosophy that implicates the use of research, evaluation, analysis, and scientific processes in law enforcement decision making. In this entry, we review the nuances of this definition, the research underpinning evidence-based policing, and what agencies employing evidence-based approaches might look like.

Fundamentals of Evidence-Based Policing

Evidence-based policing is a law-enforcement perspective and philosophy that implicates the use of research, evaluation, analysis, and