INVESTIGATION · FOUNDATIONS

Witness Statement Analysis and Consistency Checking Techniques

Investigators compare what witnesses say across interviews and against other evidence. This article explains which statement-analysis methods have scientific support, which do not, and why an inconsistent account is not necessarily a false one.

ModQu research article · 1,794 words · 18 references (APA 7th edition)

Statements are among the most common forms of evidence in any investigation. A workplace inquiry depends on what employees say happened at a meeting; a fraud review depends on how a manager explains a payment; an insurance investigation depends on a claimant's description of an accident. Investigators therefore spend much of their time comparing accounts: one witness against another, one interview against a later one, and every account against documents, data and physical evidence.

Two questions drive this work. Is the account accurate, in the sense that it matches what actually happened? And is the witness telling the truth as they remember it? The two are different. An honest witness can be mistaken, and a dishonest one can include many accurate details. This article reviews the main techniques used to analyse statements and check their consistency, and summarises what research says about how well they work.

The limits of intuition

Many people believe they can tell when someone is lying. The evidence says otherwise. In a meta-analysis of studies in which people judged whether others were lying or telling the truth, Bond and DePaulo (2006) found an average accuracy of 54%, only slightly better than the 50% expected by chance. People correctly classified 61% of truthful statements but only 47% of deceptive ones, because they tend to assume others are being honest (Figure 1).

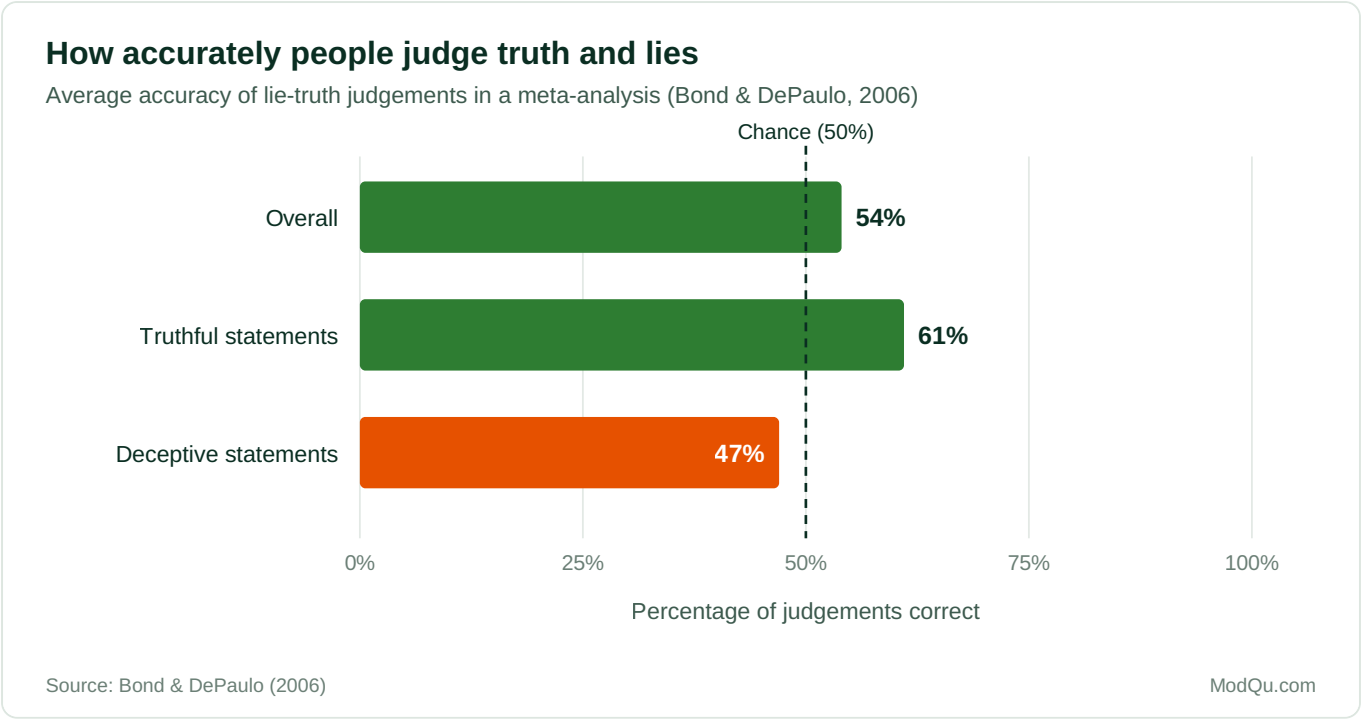


Figure 1. Average accuracy of people's judgements of whether statements are true or false, from the meta-analysis by Bond and DePaulo (2006). Chart by ModQu.

Memory research adds a second caution. Accounts can be changed by the questions asked. In a classic experiment, people who were asked how fast cars were going when they "smashed" into each other gave higher speed estimates than those asked about cars that "hit" each other, and were later more likely to report seeing broken glass that was not there (Loftus & Palmer, 1974). Decades of later work show that misleading information introduced after an event can become part of a person's memory of it (Loftus, 2005). At the same time, Wixted et al. (2018) argue that memory is more reliable than this literature is sometimes taken to imply, provided that the first account is obtained early and without contamination. The practical message for investigators is that how a statement is obtained determines how much it can later be trusted.

Obtaining a statement worth analysing

No analysis technique can rescue a statement produced by leading questions. Research-based interviewing methods are therefore the foundation of statement analysis. The **cognitive interview** uses open invitations to report everything, mental reinstatement of the context and varied retrieval cues to help cooperative witnesses recall more correct detail without a large increase in errors (Fisher & Geiselman, 1992). Statements should be recorded in full, preferably on audio or video, so that later analysis works from the witness's own words rather than an interviewer's summary.

Content-based credibility techniques

Several methods try to judge whether an account is based on real experience by examining its content.

Criteria-based content analysis

Criteria-based content analysis (CBCA) grew out of forensic work with child witnesses in Germany. It rests on the idea, often called the Undeutsch hypothesis, that accounts of experienced events differ in content and quality from invented ones. Trained assessors score a transcript against a list of 19 criteria, such as logical structure, quantity of detail, contextual embedding, descriptions of interactions, reproduction of conversations and spontaneous corrections. Vrij (2005) reviewed the first 37 studies and found that truthful accounts generally received higher CBCA scores than fabricated ones, but warned that accuracy was well short of the standard required for court evidence and that scores are affected by the witness's age, verbal skill and the interview style. A later meta-analysis also supported the hypothesis for many criteria while noting substantial variation between studies (Amado et al., 2015).

Reality monitoring

Reality monitoring draws on cognitive theory about how people distinguish memories of real events from memories of imagined ones (Johnson & Raye, 1981). Memories of experienced events tend to contain more sensory, spatial and temporal information; imagined events tend to contain more references to thoughts and reasoning. Masip et al. (2005) reviewed studies applying these criteria to deception and found accuracy rates similar to those of CBCA, again above chance but with considerable error.

Content-based techniques overall

Oberlader et al. (2016) meta-analysed content-based techniques, including CBCA and reality monitoring, and found that they discriminated between true and fabricated statements with a large average effect, but with wide differences between studies and settings. The techniques are best used as structured ways of examining an account, not as verdicts on it.

Verifiable details

A more recent approach asks a simple question: how much of the account can be checked? Nahari et al. (2014) found that truthful people tend to report more details that could be verified, such as named people, recorded transactions or locations with cameras, while liars prefer details that cannot be checked, because they want to appear cooperative without being caught. The **verifiability approach** fits naturally with investigative work, because it turns attention from the witness's manner to the evidence that can confirm or refute the account.

A method without support

Scientific Content Analysis (SCAN) is sold to police and corporate investigators as a way of detecting deception from written statements. A UK Home Office evaluation found little evidence for its claims (Smith, 2001), and an experimental test found that SCAN criteria did not distinguish truthful from fabricated accounts of a negative event (Bogaard et al., 2016). Investigators should not rely on it.

Consistency checking

Comparing repeated statements is one of the oldest investigative tools. Many investigators, lawyers and jurors treat inconsistency as a sign of lying or poor memory. The research picture is more complicated.

Types of difference

When two statements from the same witness are compared, each detail falls into one of four categories (Figure 2). It may be **consistent**, appearing in the same form both times. It may be a **reminiscence**, a detail reported later but not earlier. It may be an **omission**, reported earlier but not later. Or it may be a **contradiction**, where the two statements conflict.

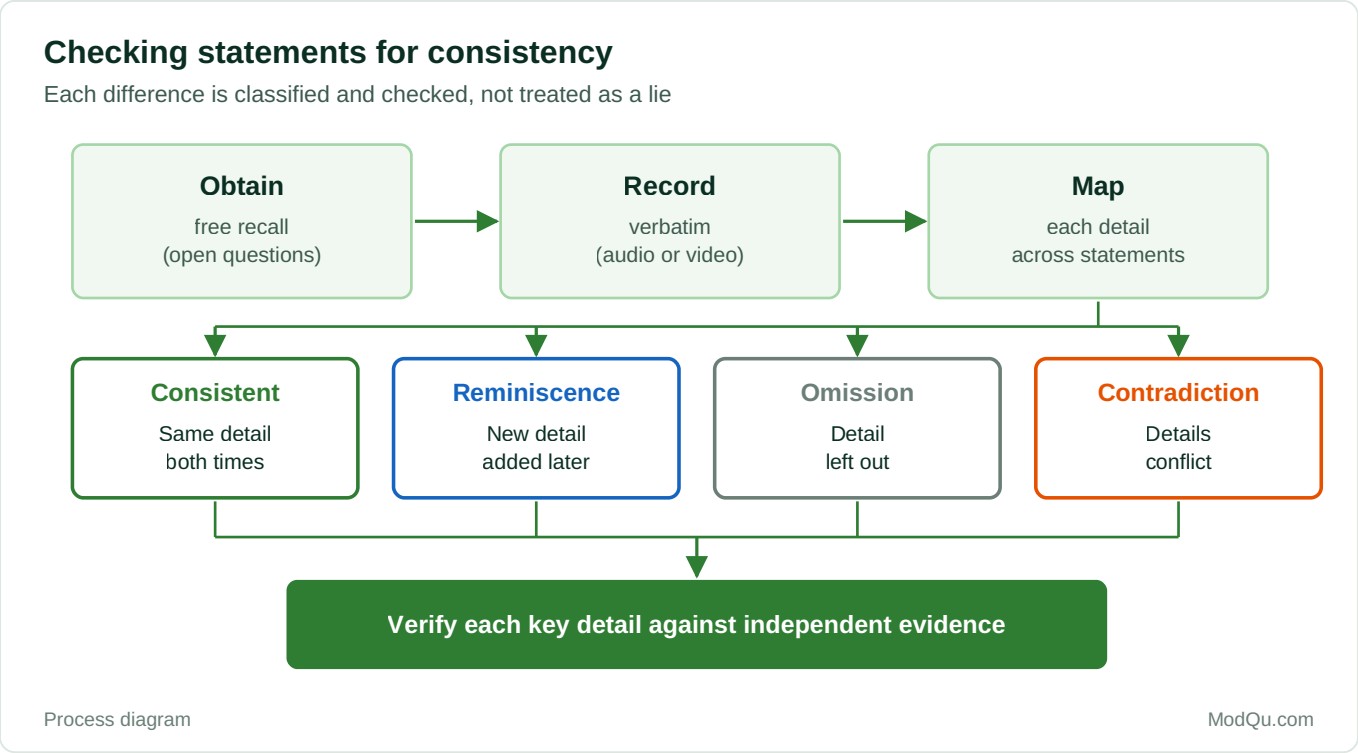


Figure 2. A structured approach to comparing statements. Differences are classified and then checked against independent evidence rather than treated as proof of deception. Process diagram by ModQu.

What the categories mean

Reminiscence is normal. Retrieval varies from one attempt to the next, so a second interview usually brings out new details. Gilbert and Fisher (2006) found that details recalled for the first time in a second interview were almost as accurate as details recalled consistently across interviews. Treating every new detail as suspicious is therefore a mistake.

Contradictions deserve more attention, because they are more often linked to errors. Even so, a contradiction shows that at least one version is wrong, not that the witness is lying.

Consistency is a weak sign of truthfulness. Granhag and Strömwall (2001) found that liars can be as consistent as truth tellers across repeated interviews, partly because liars prepare and rehearse a story and then repeat it, while truthful witnesses simply try to remember. A perfectly consistent account may be honest, or it may be a well-practised script.

Consistency with the evidence

The most useful comparison is often between a statement and independent evidence. The **strategic use of evidence** technique delays disclosing what the investigator already knows, asks the person to give an account first, and then examines whether that account contradicts the evidence. Hartwig et al. (2006) found that police trainees taught this approach detected deception far more accurately than untrained trainees. More broadly, cognitive approaches that increase the mental effort of lying, ask unexpected questions or use evidence strategically improved accuracy over standard approaches in a meta-analysis by Vrij et al. (2017).

Comparing accounts from several witnesses

Different witnesses to the same event rarely give identical accounts, and they should not be expected to. People attend to different things, stand in different places and remember different details. Small differences between independent witnesses are normal; identical wording, by contrast, can suggest that witnesses have discussed the event or been coached. Vrij (2008) notes that investigators often over-interpret both kinds of signal and recommends that conclusions rest on verifiable facts rather than on impressions of how an account sounds.

A useful practice is to build a single timeline or event matrix that lists every material detail and shows which witnesses reported it, in which interview, and whether any document or record confirms it. Details reported by several independent witnesses and supported by records form the core of what can be established. Details reported by one witness only are leads to be checked, not findings. Details on which witnesses conflict show where further enquiries are needed, for example obtaining access-card logs to settle who left a meeting first.

Witnesses should be interviewed separately and as soon as practicable, and asked not to discuss the event with one another, because shared conversation can blend memories. Where contact has already occurred, the report should say so, since agreement between witnesses who have talked is weaker evidence than agreement between witnesses who have not.

Practical guidance for investigators

1. **Get a clean first account.** Use open questions and record the answer in full before introducing any information the witness may not already have.
2. **Map details systematically.** Break each statement into discrete details and lay them side by side, rather than forming an overall impression.
3. **Classify differences before judging them.** Note reminiscences, omissions and contradictions separately, and ask the witness to explain significant contradictions.
4. **Prioritise verifiable details.** Check names, times, places, messages and transactions against records. A verified detail is worth more than any credibility score.

5. **Use content criteria as prompts, not verdicts.** CBCA and reality monitoring can show where an account is thin or rich, but they do not establish truth.
6. **Avoid unsupported tools.** Do not base findings on SCAN or on beliefs about nervous behaviour.
7. **Document the reasoning.** The investigation report should explain which details were corroborated, which were contradicted and how differences between statements were interpreted.

Key takeaways

Statement analysis is useful when it is structured, evidence-based and modest about what it can show. People judge lies only slightly better than chance, and memory can be changed by how questions are asked. Content-based methods such as CBCA, reality monitoring and the verifiability approach have research support as aids, while SCAN does not. Inconsistency is not proof of lying: new details are usually accurate, and consistency can be rehearsed. The strongest test of any statement is whether its key details agree with independent evidence.

References

Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and criteria based content analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context, 7*(1), 3–12. https://doi.org/10.1016/j.ejpal.2014.11.002

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Scientific Content Analysis (SCAN) cannot distinguish between truthful and fabricated accounts of a negative event. *Frontiers in Psychology, 7*, Article 243. https://doi.org/10.3389/fpsyg.2016.00243

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2

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

Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology, 20*(6), 723–739. https://doi.org/10.1002/acp.1232

Granhag, P. A., & Strömwall, L. A. (2001). Deception detection based on repeated interrogations. *Legal and Criminological Psychology, 6*(1), 85–101. https://doi.org/10.1348/135532501168217

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior, 30*(5), 603–619. https://doi.org/10.1007/s10979-006-9053-9

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366. https://doi.org/10.1101/lm.94705

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*(1), 99–122. https://doi.org/10.1080/10683160410001726356

Nahari, G., Vrij, A., & Fisher, R. P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology, 19*(2), 227–239. https://doi.org/10.1111/j.2044-8333.2012.02069.x

Oberlader, V. A., Naefgen, C., Koppehele-Gossel, J., Quinten, L., Banse, R., & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 40*(4), 440–457. https://doi.org/10.1037/lhb0000193

Smith, N. (2001). *Reading between the lines: An evaluation of the scientific content analysis technique (SCAN)* (Police Research Series Paper 135). Home Office.

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*(1), 3–41. https://doi.org/10.1037/1076-8971.11.1.3

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Wiley.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology, 22*(1), 1–21. https://doi.org/10.1111/lcrp.12088

Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science, 13*(3), 324–335. https://doi.org/10.1177/1745691617734878