# Lie Detection: Science or Pseudoscience?

Verschuere, B.

**DOI**
10.1007/978-3-031-75875-1_2

**Publication date**
2024

**Document Version**
Final published version

**Published in**
Legal and Forensic Psychology

CHAPTER 2

# Lie Detection: Science or Pseudoscience?

*Bruno Verschuere*

'Perhaps lie detection?' is what I muttered. It was autumn of 2000, and I was about to start a position as academic assistant at Ghent University. The position involved conducting scientific research with the aim to complete a PhD, and the topic of that research had yet to be determined. Having worked with offenders during my internship, I was aware of the pressing need for reliable ways to distinguish lies from truth. The head of department had asked me about the desired topic of my research, but my answer elicited a dismissive look and the reply 'We don't do pseudoscience in this department'. In the following 20+ years, along with many others, I have devoted my time to scientific research on lie detection (for reviews see e.g., Granhag et al., 2015; Rosenfeld, 2018), raising the question: Can there be a science of pseudoscience in the context of lie detection?

B. Verschuere (✉)
University of Amsterdam, Amsterdam, the Netherlands
e-mail: b.j.verschuere@uva.nl



I. Bošković (ed.), *Legal and Forensic Psychology*,
https://doi.org/10.1007/978-3-031-75875-1_2

## The Demarcation Problem: The Difficulty to Differentiate Science from Pseudoscience

Science and pseudoscience are not easily differentiated. For one, the topic itself or the amount of controversy around it do not allow to tell whether something is science or pseudoscience. One can conduct scientific inquiry about astrology, psychedelics in psychotherapy, or polygraphy. This problem is known in science philosophy as the demarcation problem (Letrud, 2022); see also Box 2.1. The difference between science and pseudoscience is gradual, rather than black and white. Both may involve the same topic, and even some of the same methods (e.g., conduct research and publish about it), and, yet, there are key differences between them worth pointing out. Some of the reasons being that using pseudoscientific lie methods may lead to unwarranted confidence, waste of resources, pose security risks, and to serious negative consequences on people's lives.

**Box 2.1 The Difficulty Differentiation Science from Pseudoscience: The Case of ChatGPT**

Asked whether lie detection is pseudoscience, ChatGPT 3.5 answered on January, 22, 2024, that indeed '*The use of polygraph tests for lie detection is considered by many scientists and experts to be pseudoscience*'. *ChatGPT justified its straightforward answer, by pointing to (1) Lack of a scientific basis, (2) False Positives and False Negatives, (3) Cultural and Individual Variations, and (4) Lack of Standardization. And that therefore*, '*many scientists and organizations, including the American Psychological Association, consider polygraph tests an unreliable and unscientific method for detecting deception*'.

This answer nicely illustrates that it is not easy to differentiate science from pseudoscience, and that conclusions are often based on wrong grounds. For instance, imperfection itself (i.e., the occurrence of False Positives and False Negatives errors) cannot be used to call a method pseudoscientific. All scientific instrumentation—whether it's the Hubble telescope, an x-ray, or a polygraph test—comes with an error margin. It's not the occurrence of errors itself, but not acknowledging those errors that make a test pseudoscientific.

## Warning Signs for Pseudoscience

Science tries to find the truth by systematic and objective inquiry. A key element in modern science is *falsification*, which in essence involves trying to prove yourself wrong (see also Chap. 10). Pseudoscience seeks confirmation of one's own beliefs, pretending to rely on the science. Thus, pseudoscience is a form of deception, and may be defined as 'science in disguise'. It pretends to value science and its procedures, yet doesn't fully commit to it. An example is the selective citation of studies with favourable outcomes while deliberately ignoring studies with less optimistic results. However, science and pseudoscience is not easily drafted, as scientists also sometimes engage in such bad research practices (Gøtzsche, 2022).

Many authors devoted their time to aid the discrimination between science and pseudoscience. Specifically, Lilienfeld (2010; see also Lilienfeld & Landfield, 2008) developed a list of warning signs that may point to the presence of pseudoscience. Below, I will illustrate eight warning signs of pseudoscientific practices that are currently widely employed in lie detection context, followed by an overview of the currently widely applied lie tests, see Box 2.2.

1. ***Ad hoc explanations to explain negative findings***
   Pseudoscientific claims often rely on explanations created to fit the results at hand, rather than providing an explanation that applies more generally. This way, specific excuses can protect claims from being disproved. For instance, I once asked an experienced Belgian police polygrapher how often he had erred. While having conducted hundreds of tests,[1] he could recall only a single mistake. This is remarkable given the substantial error margin of polygraph tests (Meijer & Verschuere, 2010). Furthermore, even that particular mistake was explained by the unique nature of the case, where certain questions could have triggered a traumatic experience, suggesting that this was a case-specific error, with low odds of reoccurring.
2. ***Lack of falsifiability***
   As seen above, besides being almost always applicable, pseudoscientific claims are phrased in a way that makes them hard to test. This

[1] While there exists several polygraph tests, when referring to the polygraph test I refer to the most widely used test, that is, the Control Question Polygraph Test. See also Chap. 3 and Box 2.2 in this chapter.

seems to be the main difference from scientific theories, which are falsifiable: They can be proven wrong. A common response to unfavourable findings from deception detection studies is redefining the nature of the employed tool. Scientific Content Analysis (SCAN; see Box 2.2), for instance, would not be a lie test, but mark 'hot spots'.

**Box 2.2 Currently Widely Applied Lie Tests**

There exists a great variety of lie tests. The tests that are currently most widely applied include the following:

**Behaviour detection**: Using behavioural tell signs to screen large number of people for suspicious behaviour (Weinberger, 2010). There are several variants, under different names, used for different purposes. One popular and well-known example is the 'Screening Passengers by Observation Technique (SPOT)' that was broadly rolled out after 9/11. Partly based on Ekman's work on micro-expressions, security officers were trained to detect terrorists based on 92 indicators (e.g., excessive fidgeting). It is hard to imagine how large number of people could be effectively screened on such a large number of indicators.

**SCAN**: Scientific Content Analysis (SCAN), an investigative tool taught globally to law enforcement, military, businesses, and insurance firms, is used to assess the veracity of written statements (Bogaard et al., 2016; Bogaard et al., 2021). SCAN is not based on a scientific theory but on personal insights from its developer, former Israeli police officer Avinoam Sapir. This includes highly improbably assumptions such as 'People do not lie' and 'Everyone wants to tell you everything'. These insights led to 29 SCAN criteria (e.g., that a statement should have a 20 [intro]—50 [core]—30 [outro] structure; with substantial deviations pointing to deception), with SCAN users differing in which criteria they apply.

**Voice stress analysis**: Layered Voice Analysis (LVA) is a technology that uses various vocal parameters (e.g., pitch, intonation, rhythm, and frequency modulation) to infer psychological states including dishonesty. It is widely used by US police. Despite continuous updates, empirical evidence consistently shows it performs at chance level (Eriksson & Lacerda, 2008).

**Polygraphy**: The polygraph is a device that records physiological responses—breathing, blood pressure and sweat secretion (Meijer & Verschuere, 2010). There is *no* physiological response unique to

(*continued*)

**Box 2.2 (continued)**

lying. For example, an increase in sweat secretion can come from lying, but it can also come from a thinking about something emotional. Conclusions about truthfulness, therefore, originate from comparing the response to the critical question to the responses to control questions. There are several polygraph tests and they differ in the control questions used. Most widely applied is the Control Question Technique. Its popularity lies in its broad applicability, as the critical questions are in essence a variant of 'Did you do it?' Examinees are maneuvered in answering the control questions (e.g., 'Did you ever tell an important lie?') negatively by hinting that a positive answer would make them look guilty. If the examinee responds strongly to the relevant questions, (s-)he is found deceptive. A truthful test outcome is obtained by a stronger response to the control questions. The Control Question polygraph test has been criticized on several grounds, but most importantly that it does not sufficiently protect the innocent. After all, it is evident that the relevant 'Did you do it?' question can elicit strong reactions from anyone—deceptive or truthful—simply because it is clearly the question at the heart of the matter.

Because hot spots may or may not signal deception, it becomes more difficult to prove the method wrong. A variant of this explanation is that the test is not a test at all ('It is a procedure!'), thereby evading the strict psychometric properties that psychological tests need to meet (Grubin & Madsen, 2005).

3. ***Strong reliance on anecdotal evidence***
Pseudoscience relies heavily on anecdotes rather than on the results of systematic tests to prove success. Case examples can look impressive and convincing. Even after 20 years, I still vividly remember witnessing the confession of a sex offender after being confronted with a deceptive polygraph test outcome. But was he in fact guilty? And was it the polygraph itself or the pressure of the interrogation that led to the confession? And how many guilty examinees do not

confess? Regardless of how vivid they are, anecdotes are no substitute for systematic inquiry.

4. ***Avoiding peer review***
   Peer review involves the critical evaluation of claims by relevant experts. Pseudoscience may avoid this peer review process, often by placing the burden of proof on sceptics rather than on those making the claim. Despite having the budget, companies selling lie tests often do not scientifically validate their products. This allows such companies to await study findings, then either embrace or dismiss the study based on its results, cherry-picking the ones that will help their sales.
5. ***Exaggerated claims***
   A hallmark characteristic of pseudoscience is making claims that go well beyond what the available evidence can support. Throughout history, practitioners and polygraph schools have claimed polygraph tests to be close to perfect, while scientific research has consistently pointed out a substantial error margin (Meijer & Verschuere, 2010). Another prime example is Scientific Content Analysis (SCAN) which, despite scientific evidence showing it performs at chance level in lie-truth discrimination (Bogaard et al., 2016; Bogaard et al., 2021), is claimed to be '*scientific—a consistent formula which gives consistent results. There is no guessing, or 20 steps to follow, or hours of stressful interviewing; you don't need to base an answer on the confusing signals of "body language". You need only the subject's own words, with their meaning clearly shown by SCAN. SCAN is the key to unlocking the truth*' (see https://www.lsiscan.com/intro_to_scan.htm, January 22, 2024).
6. ***Lack of self-correction***
   In science, errors are seen as highly informative. If the results do not confirm the theory, the theory may be wrong or in need of adjustment. Pseudoscientific claims may resist modification even in the face of contradictory evidence. For instance, despite scientific findings justifying questioning of its validity and evidence supporting competing tests (Bogaard et al., 2021; Oberlader et al., 2016), SCAN is still marketed as being '*the most effective technique available for obtaining information and detecting deception from statements of witnesses or suspects*', see https://www.lsiscan.com/index.htm. And the company that sells voice stress analysis, Neymesco, points to research that would underscore its deception detection potential, yet neglects to mention the key study outcome: 'Replicating earlier research the vocal analysis software's built-in deception classifier

performed at the chance level', see https://www.nemesysco.com/research/ (accessed January 22, 2024).

7. ***Use of jargon***
Also known as psychobabble or *pseudo-profound bullshit* (Pennycook et al., 2015). Pseudoscience often makes use of language that is needlessly complex, and not well-grounded in accepted scientific principles. Consider, for instance, the following polygraph theory: '*Anti-climax dampening involves the inter-relationship of two issues, questions or topics, in close proximity to each other, where the more important, bothersome or stimulating issue suppresses or completely eliminates emotional response to the other issue, question or topic which the person might have responded to had the other strong issue, question or topic not been present*' (Matte, 1996).
8. ***Appeal to authority***
Namely, the popularity and maintenance of pseudoscience usually rely on authority rather than scientific facts. In the world of security, Israel is an authoritative source. Many lie test proudly mention being developed in Israel. Another example is marketing that the lie test was used by an authoritative agency like the FBI, even though the use of an ineffective test by an authoritative agency does not make it effective. There could be several reasons other than lie detection accuracy why the test was applied. The polygraph, for instance, may be used in pre-employment screening as a deterrent and/or to invite people to admit transgressions (Meijer et al., 2008).

## Why Is Pseudoscience So Attractive?

There are a number of factors that likely contribute to the allure of pseudoscientific lie tests. Whether it is insurance companies detecting fraudulent claims, police investigators interviewing suspects, or border control officers screening for terrorists, practitioners face urgent needs that require *immediate* solutions. Scientific lie tests come with complex answers pointing to boundary conditions, and require tolerance to uncertainty (e.g., error margins). Pseudoscientific lie test, in contrast, offer straightforward answers and certainty. They are one-size-fits-all and are marketed to be applicable in almost any case. Further, there are individual as well as contextual factors contributing to the adoption of pseudoscientific lie tests (Lee & Hunsley, 2015). Individual factors include inherent biases in human thinking (e.g., 'I have seen it work, so it must be accurate'; a form of naïve realism or the tendency to consider your own perception of the world is unbiased). Contextual factors include the fact that many practitioners lack formal training in critical thinking that can counter against such biases.

## Anecdotes Versus Accuracy

An apparent success with a lie test (e.g., a confession after a deceptive test outcome) can seem very impressive. And yet, this should not be used to evaluate the accuracy of lie tests. The 'hit' should be considered in the light of the misses. Take the ADE 651, a fake bomb detector that was used at Iraqi checkpoints and was claimed to have led to the detection of several bombs. Unmentioned are the number of times that the ADE 651 warned for a bomb but there was no real danger (false positives) and the number of times it failed to detect real bomb (false negatives). This helps to explain the success of this scam https://en.wikipedia.org/wiki/ADE_651.

Rather than relying on personal experience, evaluating the accuracy of lie tests should be based on high-quality, peer-reviewed research. In defending the behaviour detection programme Screening for Passengers by Observation Techniques (SPOT), Paul Ekman argued that '*In a never publicly released study … people identified by the TSA SPOT were fifty times more likely to be wanted felons or smugglers than those selected at random*', see https://www.paulekman.com/blog/ineffectively-testing-the-effectiveness-of-tsa/. That the study is not in the open domain makes it impossible for scientist to evaluate the extraordinary claims, and the basis for it. From the brief description, it is, however, immediately clear that the control condition is problematic, and the apparent success cannot be contributed to SPOT. To explain, let us consider the typical airport situation. As the overwhelming majority of passengers do not have malicious intent, random selection is unlikely to lead to the identification of felons or smugglers. Now and then, there is someone that clearly warrants closer attention, perhaps because of being intoxicated, verbally aggressive, or carrying baby diapers while travelling solo. These people are more likely to be selected by SPOT-trained officers and to turn out to be felons or smugglers. But that is no reason to conclude '*that the system is working*', because also without being trained in SPOT, security officers are likely to have a closer look at those individuals. In sum, the inappropriate control condition of random selection inflates the estimated accuracy, and a comparison with a proper control condition (TSA officers that are as experienced and motivated as those in the SPOT condition but were not trained in SPOT) is required. Of note, SPOT was developed to detect terrorists, not smugglers and felons. Here too, caution is warranted because of post hoc cherry-picking success: If smugglers and felons are also detected, it would speak to the broad success of SPOT, if they go undetected the findings can be dismissed because SPOT was not designed to detect them in the first place.

## Towards a Science of Lie Detection

Behaviour detection programmes such as SPOT lack high-quality studies to support its effectiveness. This was also the conclusion from the Government Accountability Office: '*the available evidence does not support whether behavioral indicators, which are used in … (SPOT) program, can be used to identify persons who may pose a risk to aviation security*', see https://www.gao.gov/products/gao-14-159. Further, SCAN lack standardization and has low reliability. Initial low-quality studies reported favourable findings, but the more recent and higher-quality studies consistently show that it cannot effectively distinguish between true and false statements (Bogaard et al., 2016; Bogaard et al., 2021). For voice stress analysis (in particular layered voice analysis or LVA), study findings are remarkably consistent in showing that one might as well toss a coin to decide whether someone is lying or telling the truth, as it does not perform better than chance (Eriksson & Lacerda, 2008). Polygraphy is the only popular lie test for which there is a substantial data base. And, although there is no agreement on the precise error margin, findings allow to conclude that (1) the accuracy of the polygraph is above the chance level, and (2) errors often occur, particularly falsely accusing truth teller to lie (Meijer & Verschuere, 2010).

Applying tests with unknown validity or with a known imperfect validity may be a waste of resources, but doing so in itself is not pseudoscience. What would be pseudoscientific is to make accuracy claims that are not supported by the evidence. In a recent murder case in Belgium, a police polygrapher communicated to the examinee that '*There is no doubt in my mind, no doubt at all—You show deceptive responses*', see https://www.hln.be/alken/assisen-loog-de-beschuldigde-of-niet-gefaalde-leugendetectortest-leidt-tot-discussie-op-proces-tiramisumoord~a37f68f7/. And in the subsequent court case, the police polygrapher argued that the accuracy of the employed lie test is 92–97%. These claims do not fit with the literature nor with expert opinion (Iacono & Lykken, 1997). Striving for scientific use of the polygraph would first require accurate communication about polygraph accuracy. It is understandable that concluding that the accuracy of polygraph tests is 'better than chance, yet well below perfection' may sound unsatisfying, yet it needs to be recognized that this level of precision is the only one that the currently available data allow for (Meijer & Verschuere, 2010).

## Currently Applied Lie Tests: Science or Pseudoscientific?

One way to evaluate the pseudoscientific nature of lie tests is to evaluate their current practice with Lilienfeld's 2010 warning sign list. The more signs are met, the more the test shifts towards the pseudoscience side of the science-pseudoscience continuum. From the examples above, it is clear that, to various degree, the most widely applied lie tests are all practiced pseudoscientifically.

Inquiring relevant experts may also be of help to scientific nature of lie tests. Iacono and Lykken (1997) surveyed 183 scientists on their opinion of the polygraph. The majority of these scientists were of the opinion that polygraph test does not have a sound theoretical foundation, that claim of high accuracy (>85%) are unfounded, that polygraph test can be circumvented by easily learned tricks, and that its test result should not be admissible in court as evidence. More recently, Bogaard et al. (2021) invited experts to express their opinion on SCAN as a lie detection tool. Researchers who had recently published at least two peer-reviewed on verbal lie detection, or at least one on SCAN, specifically, were invited. When asked to indicate the scientific status of SCAN on a scale of −100 (pseudoscientific) to +100 (scientific), researchers obtained the mean score of −59 (SD = 41). Rescaled to a binary decision, 88% of experts qualified SCAN to be pseudoscientific. As far as I know, there have been no surveys yet to poll scientists on their evaluation of behaviour detection programmes like SPOT or on voice stress analysis, which would be a valuable insight when discussing the current state of art of lie detection.

## Conclusion

Lie tests are the psychological tests with the greatest possible impact for an individual, urging the need for scientists to critically study them (Lykken, 1974). Such need is also supported by the public, who also desires scientists to be involved (Cologna et al., 2024). Science of lie detection is of key interest also for policy makers, who should be able to rely on scientific evidence when considering the use of lie tests. The most widely applied lie tests are currently practiced in a pseudoscientific manner. The conclusion of this chapter is not as pessimistic as it may now appear. Lie tests are not necessarily pseudoscientific: Even when being far from perfect, they can be practiced in a scientific manner. The most important step forward is to realize that 'Doubt is not a pleasant condition, but certainty is absurd'.[2]

## References

Bogaard, G., Verschuere, B., & Meijer, E. (2021). Centre for policing and security: Stop met het aanbieden van pseudowetenschap. *Panopticon, 42*(4), 359–361.

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*. https://doi.org/10.3389/fpsyg.2016.00243

Cologna, V., Mede, N. G., Besley, J. C., & Maibach, E. W. (2024). Trust in scientists and their role in society across 67 countries Trust in scientists and their role in society across 67 countries. *OSF Preprint*. https://osf.io/preprints/osf/6ay7s

Eriksson, A., & Lacerda, F. (2008). Charlantry in forensic speech science: A problem to be taken seriously. *International Journal of Speech, Language and the Law, 14*(2), 169–193. https://doi.org/10.1558/ijsll.v14i2.169

Gøtzsche, P. C. (2022). Citation bias: Questionable research practice or scientific misconduct? *Journal of the Royal Society of Medicine, 115*(1), 31–35. https://doi.org/10.1177/01410768221075881

Granhag, P. A., Vrij, A., & Verschuere, B. (2015). *Detecting deception: Current challenges and cognitive approaches*. Wiley-Blackwell.

Grubin, D., & Madsen, L. (2005). Lie detection and the polygraph: A historical review. *Journal of Forensic Psychiatry and Psychology, 16*(2), 357–369. https://doi.org/10.1080/14789940412331337353

[2] Quote commonly attributed to the French philosopher Voltaire.

Iacono, W. G., & Lykken, D. T. (1997). The validity of the Lie detector: Two surveys of scientific opinion. *Journal of Applied Psychology, 82*(*3*), 426–433. https://doi.org/10.1037/0021-9010.82.3.426

Lee, C. M., & Hunsley, J. (2015). Evidence-based practice: Separating science from pseudoscience. *Canadian Journal of Psychiatry, 60*(*12*), 534–540. https://doi.org/10.1177/070674371506001203

Letrud, K. (2022). Incorrigible science and doctrinal pseudoscience. *International Studies in the Philosophy of Science, 35*(*3–4*), 269–278. https://doi.org/10.1080/02698595.2023.2218264

Lilienfeld, S. O. (2010). Can psychology become a science? *Personality and Individual Differences, 49*(*4*), 281–288. https://doi.org/10.1016/j.paid.2010.01.024

Lilienfeld, S. O., & Landfield, K. (2008). Science and pseudoscience in law enforcement: A user-friendly primer. *Criminal Justice and Behavior, 35*(*10*), 1215–1230. https://doi.org/10.1177/0093854808321526

Lykken, D. T. (1974). Psychology and the lie detector industry. *The American Psychologist, 29*(*10*), 725–739. https://doi.org/10.1037/h0037441

Matte, J. A. (1996). *Forensic psychophysiology using the polygraph: Scientific truth verification; Lie Detection*. J.A.M. Publications Address.

Meijer, E. H., & Verschuere, B. (2010). The polygraph and the detection of deception. *Journal of Forensic Psychology Practice, 10*(*4*). https://doi.org/10.1080/15228932.2010.481237

Meijer, E. H., Verschuere, B., Merckelbach, H. L. G. J., & Crombez, G. (2008). Sex offender management using the polygraph: A critical review. *International Journal of Law and Psychiatry, 31*(*5*). https://doi.org/10.1016/j.ijlp.2008.08.007

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 & Human Behavior, 40*(*4*), 440–457.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making, 10*(*6*), 549–563. https://doi.org/10.1017/s1930297500006999

Rosenfeld, J. P. (2018). *Detecting concealed information and deception: Recent developments*. Elsevier Academic Press.

Weinberger, S. (2010). Airport security: Intent to deceive? *Nature, 465*, 412–415. https://doi.org/10.1038/465412a