

# Scientific Content Analysis (Scan): Perceptions about the Technique Dedicated to Detecting Deception.

Letícia Barros Corrêa de Lima

Department of Public Safety, Fluminense Federal University (UFF), Niterói, Brazil, leticiacorrea@id.uff.br.

**Abstract.** SCAN stands out for Scientific Content Analysis, a technique developed with the aim of revealing information and detecting if deception is present within written statements. Throughout the years, SCAN has proved its importance by being vastly used in continents such as North America, Australia, and Europe and in an expressive number of services of national and global importance dedicated to law enforcement, military intelligence, and federal agencies. This paper was established with the purpose of comprehending how this statement analysis technique is expected to perform in practice, what is offered to those responsible for interpreting written sentences, different ways of approaching Scientific Content Analysis, as well as who is currently invested in learning this tool. The constant criticism attributed to SCAN due to its supposed lack of scientific foundation was also discussed here. As the study was being conducted, it became clear that the analysis of SCAN reports and impressions of previous researchers provided lenses to understand challenges offered given the insufficient detailment of ways available for validating criteria and the absence of a theoretical support for the technique. SCAN ended up demonstrating that it is possible that previous experiences of trained professionals might influence how they analyse a statement, suggesting that the technique could have been applied in a non-standardised format, even with the help of over 10 criteria. It also indicated that SCAN could benefit from the development of a theory that supported sociocultural features and different backgrounds capable of influencing one's organisation of sentences when providing a statement.

**Keywords.** Scientific Content Analysis, Statement Analysis, Content-Based Techniques, Investigation, Police Interview

## 1. Introduction

The idea of having a technique that serves as a tool capable of identifying differences between truth and deception is highly appealing to most of us for uncountable reasons. Now, imagine the impact this sort of ability is entitled to have in legal contexts such as statements given by witnesses, sources or even suspects. This is what Scientific Content Analysis — mostly known as SCAN — and many other techniques promise to offer to thousands of detectives all around the world.

Created by the former polygraph examiner and lieutenant Avinoam Sapir, SCAN is promised to easily distinguish truth from deceit through an analysis of written statements with the help of 13 criteria (1) established within the technique, acknowledging that the number of criteria differs according to the study — some may show the presence of 12 criteria. Training is given by Sapir itself, whether in-person or not since an online form of studies were also proposed due to the pandemic of COVID-19. Even though further research needs to be conducted in

order to fully comprehend the level of effectiveness, the Scientific Content Analysis technique is widely spread worldwide and in all types of agencies out there. When entering the official website dedicated to comprising information about SCAN, the Laboratory for Scientific Interrogation (LSI) (2) it is possible to see a considerable number of places that have already received training in this deception-identifier tool. Canada, Mexico, and the United States are just a few examples.

Relying on the interest to understand what is this tool so commonly used in investigations, this paper is committed to discussing topics regarding the way SCAN is promised to perform technically, who is investing in it and the usual concerns of researchers relating to whether a scientific approach is present within the Scientific Content Analysis technique. Notwithstanding, the LSI website is fundamental for comprehending what we should expect from SCAN, and in spite of not being extensively mentioned in projects dedicated to analysing the subject, it appears as an important source for this present work.

## 2. Research Methods

This research was conducted with the help of the existent literature in the subject, therefore through bibliographic research. A reasonable number of articles and books were gathered during the period of the scientific training provided by the Institute of Czech-Brazilian Academic Cooperation at the UNIGOU Program. Nonetheless, it was challenging to collect data in a reality of a considerable amount of papers not being available for online public access and in a context of an already restricted field of study. Still, the present work proceeded to overcome its challenges collecting valid accessible scientific projects relating to the subject and comprehending the perspective of SCAN users with an aim to identify if the current experience with the technique has been aligned with what Scientific Content Analysis is expected to be.

## 3. SCAN in practice

Scientific Content Analysis has the main purpose of identifying differences between a truthful and a deceitful written statement through the own words of those investigated, looking out for concealed lies and other facts that are being disguised (1). For this to happen, it is asked for the individual, whether a suspect or not, to provide his/ her version of what happened without being influenced by the investigator by any means. The statement will be later analysed by a trained professional in SCAN, without necessarily being the one that was present at the moment it was written, since Sapir's technique is sold with the promise of qualifying individuals for precise analyses without any knowledge of the person investigated, nor the case itself. Some criteria present within the SCAN technique are more qualified to be used for the direct identification of deception or truth, while others may be used interchangeably (3). SCAN is also said to be of worth for any professional whose work depends on collecting information from written material, as it can be seen in the home page at the LSI website.

The criteria are known as '*spontaneous corrections*', '*first person singular past tense*', '*location of emotions*', '*change in language*', '*lack of conviction and memory*', '*extraneous information*', '*missing information*', '*out of sequence information*', '*objective and subjective time*', '*social introduction*', '*pronouns*', '*structure of statement*' and '*denial of allegation*'. Therefore, SCAN criteria are supposed to help professionals to analyse the organisation of speech of those who are providing a statement as it takes conclusions on significations based on the structure of phrases and construction of ideas. (1). Studies have shown that '*change in language*' is one of the most common criteria pointed out by SCAN users as an effective indicator of deception (4) (5). This would stand for situations where an individual is describing an event and suddenly starts to use different words to talk about the same item or situation, thereby seen as a major inconsistency on someone's statement. It

implicates a change in reality. One research demonstrates that this criterion was the only one that accurately distinguished truth from deceit, which implies that further research should be conducted to study this phenomenon (4).

Still talking about how SCAN trained professionals validate statements, it is vital to emphasise the meaning attributed to the '*lack of conviction and memory*' criterion. In Scientific Content Analysis, hesitation and uncertainty are seen as red flags for deception. Meanwhile, when we look at other techniques used in investigative settings such as Criteria-Based Content Analysis (CBCA), this act is interpreted as a sign of authenticity (3) (4). Not to mention that it is common that individuals that have gone through stressful experiences, as expected for those who are being analysed under SCAN, tend to face some level of trouble when remembering certain occurrences (6). Besides, giving a statement, does not matter if in a written or oral form, is a situation that puts individuals under severe pressure. However, in a broad scope, SCAN does not appear to perform above luck due to its low rates of accuracy, as well as the alike performance of experienced detectives who were not even trained in the technique, implying that Scientific Content Analysis does not do much different from what is already there, a hypothesis that still needs to be properly explored (5).

Despite focusing its analytical efforts on written statements, SCAN can also be encountered in a questionnaire format, which is called Verbal Inquiry - the Effective Witness (VIEW). The VIEW questionnaire, as mentioned by its creator, is a 'ready-to-use' form, therefore, able to save time of the interviewer by sending it through fax or email to those involved in the case without experiencing geographic barriers. Nonetheless, it is difficult to find research that furnishes an extensive look into this other way of approaching the SCAN technique. There is not much to offer in the LSI website as well; VIEW is simply described by Sapir in the LSI's website (7) as a practical and effective tool that will make answers "jump out at you" and that will "practically solve the case by itself", but no support is given on how the method actually works.

Throughout the years, a considerable amount of research was conducted with the appeal of identifying if SCAN could be rightly classified as scientific or not. The lack of a theoretical foundation, for instance, hardens the comprehension of how it is done, and which situations are applicable for using the technique (8). This not only makes it difficult to attest how trustable the technique is, but also indicates that professionals trained in SCAN might end up using the tool differently, attributing different meanings for the criteria and applying them in different occasions. Furthermore, research has shown that there is not a fixed list of criteria to be used, meaning that it is not possible to say that all analyses are performed equally, as well as it remains unclear why certain criteria are not considered in

some statements (4). Even though there are descriptions available of SCAN criteria, it is hard to understand why certain choices of words and types of organising one's speech are seen as signals of deception or attempts to disguise involvement in the occurrence.

#### 4. Who is funding SCAN?

Knowing that SCAN is considerably used by several countries, it comes to mind which agencies have been responsible for the interest and the investment of training professionals in this technique. In the Laboratory for Scientific Interrogation website, it is possible to encounter a list (9) of military law enforcement and intelligence, corrections, insurance companies and federal agencies that already participated in the SCAN training, including the National Security Agency (NSA), the Federal Bureau of Investigation (FBI) and the Central Intelligence Agency (CIA).

In 2016, a study (10) was conducted by the High-Value Detainee Interrogation Group (HIG), an agency administered by the Federal Bureau of Investigation (FBI) and with the participation of professionals from the Central Intelligence Agency (CIA) and the Department of Defense (DoD), to analyse the effectiveness of interrogation techniques. In their SCAN analysis it was concluded that "SCAN is widely employed in spite of a lack of supporting research" and that some criteria used to interpret statements are also used by other investigative techniques, however acting out as different indicators, something that was already noticeable throughout the reading of previous studies as it was mentioned. Even though it has the collaboration of US federal institutions that have recognised, in their own words, that "the efficacy of the technique is difficult to assert", Scientific Content Analysis continues to be extensively used in the country and with the known support of incredibly empowered agencies that have a major global impact.

#### 5. How Sapir displays the technique through case studies

There is no better way to fully comprehend the use of a technique than observing how it works in practice with its own creator. In the Laboratory for Scientific Interrogation website, there is a section thoroughly dedicated to SCAN reports where Sapir provides extensive analyses on 16 reports of the most varied topics, including politics and criminal cases of national / international repercussions (11). The reports are divided in 4 sections: The World Trade Center, Congressman Gary Condit, The Clinton Affair and DC Politics and Old Cases.

Sapir's analysis on the JonBenet Ramsey case allows us to understand how Scientific Content Analysis criteria are expected to be used and which results we

can await from them. By extracting sentences out of an interview given by JonBenet's parents, John and Patsy Ramsey, Sapir has the intention of finding clues on what actually happened with the 6-year-old superstar whose murder remains a mystery to this day. Here, I would like to shed light on the analysis performed with John Ramsey sentences since this will be beneficial to understand how Avinoam Sapir achieved his final conclusions on the case. John constantly refers to his daughter by her own name or simply as a 'child' anytime he mentions how JonBenet used to be when she was alive. However, when talking about her demise, John refers to her as a 'daughter'. Sapir says that it is not unusual to see change in language when talking about children who were abused and ended-up being killed by their perpetrators. It is important to remember that '*change in language*' is one of the 13 criteria present within SCAN. Besides, it is said to be expected that those parents who abuse their child experience some trouble referring to them directly as 'son' or 'daughter', since it causes some level of discomfort for the speaker. Sapir also notices that throughout the interview John Ramsey constantly violates the '*first person singular past tense*' criterion, which, according to SCAN, is dedicated to identifying an individual's commitment with its own sentences. At this point, John's words are looking deceitful through Scientific Content Analysis lenses and implicate that he is involved in JonBenet's murder.

It is also not an easy task to find content of Mr. Sapir himself discussing with others his own creation, the Scientific Content Analysis technique, available for free on the internet. In 2019, an investigative report was conducted by ProPublica (12) in order to understand what SCAN is and why it has been so massively used all around the world. The report presented exclusive excerpts from an interview given by Avinoam Sapir in a local television program a few years ago and that was acquired by the investigative journal through a DVD purchased at the LSI website.

In this case, Sapir is talking about Anita Hill, a former worker of the US Department of Education's Office for Civil Rights. There, she worked as a legal advisor for Clarence Thomas, who was an assistant secretary of the department and a Supreme Court nominee by the time of the occurrence (1991). Hill claimed that Clarence Thomas assaulted her sexually, and this is what Sapir will analyse in the interview. Sapir then draws attention to the fact that, during her testimony, Anita Hill only referred to herself as an 'individual' and a 'person', but never as a 'woman'. This fact by itself would be enough to imply that Anita was, in fact, someone who struggled with her sexual identity. At a certain moment, the former OCR worker also mentions that she "had a normal life with other men outside the office", which Sapir mentions in the interview as a clue that Anita was once labelled as abnormal before since this would be the only reason for someone to call themselves as 'normal'.

Both situations described in this present section reflect an intense level of specificity required to apply the Scientific Content Analysis technique. When watching how SCAN performs on real cases, it is common to see Mr Sapir focusing on small fragments of talk, trying to look out for patterns of normality expected within one's speech, highlighting that 'normality' here is mentioned in the sense of the expectancy in identifying certain elements of truth. It is true that criteria used in the Scientific Content Analysis technique analyses ways of transmitting information, patterns, and the structure present in speeches. Yet, it is noticeable that SCAN is also concerned with certain choices of words, regardless of the 'change in language' criterion. Alike Anita Hill's case where it was implied that she most likely struggled with her sexuality solely based on her preference for using the word 'person' instead of 'woman' when referring to herself, other analyses provided by Avinoam Sapir also indicate the presence of worries with such details. The use of the word 'child' in sentences can be seen as a possible symbolism of sexual abuse being experienced previously, as well as a constant mentioning of the opening and closing of a door could also imply the experience of childhood abuse. There is not a criterion especially dedicated to isolated words or sentences that could represent metaphors and implicit indications of events, nor a theory established to understand such phenomenon, even though it seems to be relatively important for SCAN due to its prevalence. This means that further research should be conducted to understand if this is also common in other SCAN-trained professionals' analyses.

## 6. Conclusions

Knowing the value of considering sociocultural differences when developing a methodology (13), it is also important to mention that the way we speak reflects who we are and from where we come from. When analysing how a speech is organised, considering social factors, historical context and personal background is a must in order to comprehend which kind of organisation of talk can be expected from a certain individual. Not ignoring the context of a criminal investigation, thereby understanding the limits of dedicating time for such analyses, I am motivated by the idea that these questions should at least provoke some reflections when it comes the time of providing a strong theoretical foundation for Scientific Content Analysis. In summary, all the discussion proposed by researchers on the SCAN technique is a result of the worldwide popularity of it. The topics brought up in such studies are crucial to improving the quality of content-based analysis made with the criteria proposed by Avinoam Sapir. Comprehending when, how and with whom SCAN should be used is necessary not only for achieving a solid performance of those trained in the technique, but also for the sake of justice.

Despite acknowledging that everyone has their own language code and idiosyncratic way of communicating, SCAN appears to draw too much attention on certain words used and their placement in sentences without taking into consideration a much broader perspective on cases. When this is combined with an inconsistent support of theories, it allows individuals to see the Scientific Content Analysis technique as deliberative on some level, according more with particular professional experiences of those who are in charge of analysing the written statements and less with the criteria previously established. It becomes difficult to understand why such affirmatives are being made, which endorsements are involved in it and if there is also a guarantee that differences between interpretations proposed by SCAN-trained analysts are not going to be too expressive.

Relating to other topics that could also benefit from research, it is known that acquiring training in this deception-identifier tool is said to be effective for anyone who deals with written statements on a daily basis, independent of which field of work one comes from. However, as we navigate through research studies conducted on SCAN, it is possible to see that most of them are dedicated to studying the usage of this technique by police enforcement. Furthermore, the Laboratory for Scientific Interrogation website appeared in this project as a meaningful source for understanding how Scientific Content Analysis is meant to be interpreted and applied by its users through the words of the creator of the technique itself, despite being mentioned only in few researches. It is a matter of fact that the website still lacks the ability to supply all the needed information regarding SCAN, but, on the other hand, it is capable of offering an essential perspective on which kind of material SCAN-trained professionals receive, different ways of applying SCAN, as well as who is currently interested in investing in it. The highlight of it, however, is with the SCAN reports available to read in a section thoroughly dedicated to this topic. For those reasons solely, I see the LSI website as an effective tool and an interesting source to be analysed in future projects dedicated to the subject.

The number of studies committed to studying SCAN also remains unsatisfactory. For this reason, it is valid to remember how widely used this technique is all around the world. This, in addition to the scientific support yet to be proven for Scientific Content Analysis, should only stimulate researchers to continually investigate how this tool works, how it has been used, who is interested in investing in it, what are the future approaches and many other topics relating to the application of SCAN. Comparative studies with other types of content-based techniques are also of great help when developing research questions regarding Mr. Sapir's invention.

## 7. Acknowledgments

This paperwork did not receive funding, nor has benefited from the participation of other people.

## 8. References

(1) Nahari G., Vrij A., Fisher R. Scientific Content Analysis (SCAN) Cannot Distinguish Between Truthful and Fabricated Accounts of a Negative Event. *Law and Human Behavior*. 2012; 36(1): 68-76.

(2) Laboratory for Scientific Interrogation [cited 2022 Apr 01]. Available from: <http://www.lsiscan.com/>

(3) Oberlader V., Quinten L., Banse R., et al. Validity of content-based techniques for credibility assessment — How telling is an extended meta-analysis taking research bias into account? *Applied Cognitive Psychology*. 2021; 35: 393-410.

(4) Bogaard G., Vrij A., Meijer E., et al. Scientific Content Analysis (SCAN) Cannot Distinguish Between Truthful and Fabricated Accounts of a Negative Event. *Frontiers in Psychology*. 7(243): 1-7.

(5) Smith N. *Reading between the lines: An evaluation of the scientific content analysis technique (SCAN)*. UK Home Office, London; 2001; 42 p.

(6) Schwabe L. Memory under stress: from single systems to network changes. *European Journal of Neuroscience*. 2017; 45: 478-489.

(7) Laboratory for Scientific Interrogation [cited 2022 Apr 01]. Available from: [http://www.lsiscan.com/intro\\_to\\_scan.htm](http://www.lsiscan.com/intro_to_scan.htm)

(8) Vrij A. *Verbal Lie Detection Tools: Statement Validity Analysis, Reality Monitoring and Scientific Content Analysis*. In: Granhag P., Vrij A., Verschuere B. *Detecting Deception: Current Challenges and Cognitive Approaches*. West Sussex: Wiley Blackwell. 2015; 3-36

(9) Laboratory for Scientific Interrogation [cited 2022 Mar 22]. Available from: <http://www.lsiscan.com/id29.htm>

(10) High-Value Detainee Interrogation Group [cited 2022 Apr 02]. Available from: <https://www.fbi.gov/file-repository/hig-report-interrogation-a-review-of-the-science-september-2016.pdf/view>

(11) Laboratory for Scientific Interrogation [cited 2022 Apr 01]. Available from: <http://www.lsiscan.com/reports.htm>

(12) ProPublica [cited 2022 Apr 02]. Available from: <https://www.propublica.org/article/why-are-cops-around-the-world-using-this-outlandish-mindreading-tool>

(13) Shuy R. *The language of confession, interrogation, and deception*. Sage, Thousand Oaks; 1998. 205 p.