

# When a Lie is More Believable than the Truth: The Dynamics of Lying and Discourse Analysis

Cuando una mentira es más creíble que la verdad: la dinámica  
de la mentira y el análisis del discurso

Erwin Warkentin, PhD  
Memorial University, Canada

e-mail: ejwarken@mun.ca

Recibido: 12/12/2020  
Aprobado: 02/02/2021

## ABSTRACT

This article describes the preliminary findings of a research project that is investigating whether there are stable and consistent personality traits found in propaganda texts that indicate whether a statement disseminated via a media is truthful or not. It is not often the case that official government sources will clearly state whether a statement previously release was intended to be truthful or not. Even when a statement is shown to be false, the claim is usually made that it was unintentionally wrong or misleading. The project uses the Rumours Broadcasts targeting German troops and civilians in France, created by the British Political Warfare Executive from July 1942 to May 1945, consisting of over 200,000 words, to analyse the personality differences between true and false statements made in those broadcasts. The first question answered in the study is the amount of fictional versus non-fictional material is necessary to make a broadcast believable. This establishes a benchmark for determining consistent differences between true and false statements in a news release or broadcasts. The analysis is done using tools developed by IBM to examine the vast amounts of data created and displayed on various social media platforms in accordance with the Big Five Personality Traits Theory. While most of the personality traits identified by these tools do not reveal any significant differences, there are some, such as a lack of imagination, cautiousness, a lack of willingness to compromise, and worry, that show consistent significant differences between the fictional and non-fictional statements.

*Keywords:* big data, lying, misinformation, Big Five Personality Traits, truthfulness, Political Warfare Executive, propaganda, social media, truthfulness, Watson, World War II.

## RESUMEN

Este artículo describe los hallazgos preliminares de un proyecto de investigación que estudia si se encuentran rasgos de personalidad estables y consistentes en los textos de propaganda que indican si una declaración difundida a través de los medios de comunicación es veraz o no. No es frecuente que las fuentes oficiales del gobierno indiquen claramente si una declaración publicada anteriormente tenía la intención de ser veraz o no. Incluso cuando se demuestra que una declaración es falsa, generalmente se afirma que fue involuntariamente incorrecta o engañosa. El proyecto utiliza las transmisiones de rumores dirigidas a tropas alemanas y civiles en Francia, creadas por el Ejecutivo de Guerra Política Británica de julio de 1942 a mayo de 1945, que consta de más de 200,000 palabras, para analizar las diferencias de personalidad entre las declaraciones verdaderas y falsas hechas en esas transmisiones. La primera pregunta que se responde en el estudio es la cantidad de material ficticio versus no ficticio que se necesita para que una transmisión sea creíble. Esto establece un punto de referencia para determinar diferencias consistentes entre declaraciones verdaderas y falsas en un comunicado de prensa o transmisiones. El análisis se realiza utilizando herramientas desarrolladas por IBM para examinar la gran cantidad de datos creados y mostrados en varias plataformas de redes sociales de acuerdo con la Teoría de los Cinco Grandes Rasgos de Personalidad. Si bien la mayoría de los rasgos de personalidad identificados por estas herramientas no revelan diferencias significativas, hay algunos, como la falta de imaginación, cautela, falta de

voluntad para comprometerse y preocupación, que muestran diferencias significativas consistentes entre lo ficticio y lo no ficticio enunciados de ficción.

*Palabras Claves:* grandes datos, mentira, desinformación, los cinco grandes rasgos de personalidad, veracidad, ejecutivo de guerra política, propaganda, redes sociales, veracidad, Watson, Segunda Guerra Mundial.

## INTRODUCTION

We are all implicitly and explicitly taught as a first principle when we learn any language, whether it is our mother tongue or our four additional languages, that we should tell the truth. We are rewarded for truth-telling and censured for lying. However, social pressures being what they are; the truth is not always what we might want to convey. We may, at times, wish to massage the truth or employ a convenient lie. Such behaviours can be challenging in our mother tongue and even more difficult in a second or third language. This quickly reveals itself in the language classroom. So those students do not reveal information that they do not want others to be privy to (home addresses, phone numbers, ages etc.) students are sometimes told that they do not need to be 100% honest with the instructor when asked to speak about them in an oral exercise in the language classroom. Sometimes, students simply do not wish to divulge this information to other class members for a variety of personal reasons. The important thing is that the teacher knows who they are. Moreover, the instructor is only interested in the student's ability to learn the target language and navigate its many phonological and grammatical difficulties.

Over the many years of teaching German at the university level, it has become quite clear to me that lying is a complicated task to accomplish in a foreign language. The semantic gymnastics required to maintain the integrity of a lie are taxing for a native speaker and even more so for a non-native speaker. This is true for a variety of different reasons. However, there are some individuals and, more importantly, organisations that are quite expert at lying. They may even take particular professional pride in being able to formulate a good lie because their job is to develop believable truths in the face of an enemy that is expecting to hear lies emanating from the other side.

While conducting other unrelated research into British World War Two propaganda broadcasts to Germany, it became clear that the British clandestine organisations tasked with delivering propaganda to Axis nations, with a particular focus on Germany, faced the same difficulty. They needed to be able to sew discord and discontent in Germany and the countries it occupied in a believable but, of necessity, untrue manner. To accomplish this, the linguists, sociologists, political scientists, journalists and Germanists employed by Great Britain's ultra-secret Political Warfare Executive (PWE) taught their propagandists and political warriors how to formulate a believable lie.<sup>1</sup>

The above revealed that more than one academic discipline would need to be brought to bear on the problem. As a result of bringing these seemingly unrelated areas of research together, it became clear that one could help clarify questions raised by the others.

The Political Warfare Executive went to great pains to ensure that their broadcasts to Germany were believable. They seem to imagine that they could make a broadcast so credible that their staff might have difficulty determining what was true and what was not. As a result of this, the authors carefully marked up their broadcast scripts, totaling approximately 207,000 words sent out over the airwaves between July 1943 and March 1944, indicating precisely what in their broadcasts was true and what was false. As a result of their careful "side lining," one can separate the text into truthful statements and false statements. Following that, it is possible

---

<sup>1</sup> The surviving records of the Political Warfare Executive are on deposit at the National Archive (Kew) in the United Kingdom and can be referenced under FO 898.

to use various text analysis or text mining programmes to identify possible cues or "tells"<sup>2</sup> as to what happens when one begins to lie in a foreign language. In this paper, it will be shown how the use of adjectives, adverbs, sentence length, and other linguistic "tells" can help determine the veracity of a given statement. Additionally, the initial findings of this project may eventually lead to a better understanding of data-mining techniques that purport to identify everything from emotional states to whom one is most likely to vote for in a coming election or what might be presented to them as a five-second video ad at the beginning of a YouTube video.

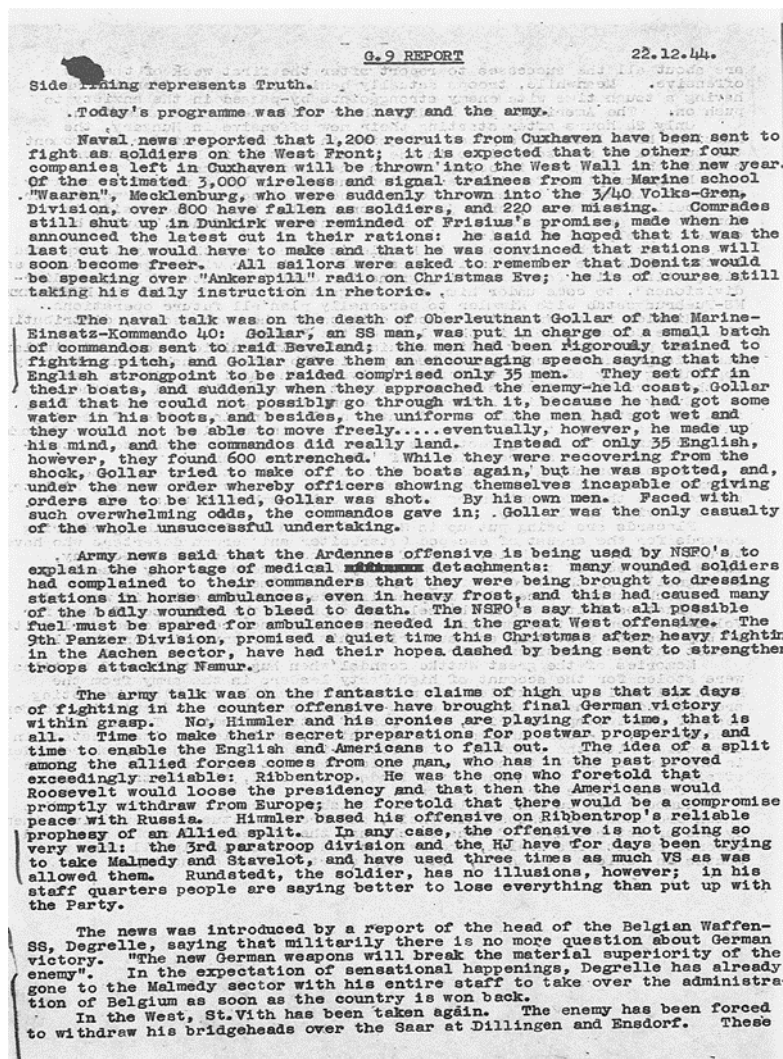


Figure 1 Example of text and "side lining" FO 898-72

These preliminary findings will focus on an area within the more considerable discipline of applied linguistics, which, for a long time, had remained on its fringes and was of interest mostly to those who investigated police interrogation techniques or were working in the legal profession and were trying to establish how one might determine whether a witness was lying or telling the truth. Of course, there had always been an interest in the dark world of domestic and international political warfare (or propaganda), but that was still considered as the illegitimate child of the applied linguistics that was never really acknowledged by the parents that had given it life.

<sup>2</sup> A "tell," when used as a noun, is an unconscious action that is thought to betray an attempt at deception.

However, since the advent of Donald Trump and the misinformation campaigns of the 2016 and 2020 elections (Russian and otherwise) as well as Boris Johnson and Nigel Farage and the Brexit vote in the UK, it has once more garnered considerable interest. Due to the disinformation and confusion that is being spread via social media outlets by various governmental and non-governmental actors, it has achieved the greater goal of causing people to mistrust established news sources. While scholars and private companies furiously mine the data of these sources, most of which is simply dross and slag, they have largely forgotten that what we are experiencing today is not without precedent. There had already been one earlier attempt at spreading misinformation by an electronic medium that was systematically planned and executed almost 80 years ago in the darkest days of the Second World War. It is of interest today in that it provides us with data on how misinformation campaigns are conceived and executed. Most importantly, it tells us precisely when a lie is being told and when the propagandist is telling the truth. These are the script left behind by the PWE (Political Warfare Executive) that was broadcasting black propaganda to the German troops in the coastal regions of France, who served both as occupiers and as personnel manning the U-boats attempted to starve Britain into submission.

It is not easy to guess what is true and what is not. The reason for this is that people are generally unwilling to reveal that they have lied. The two leading reasons are that there is a stigma attached to being a liar and that one would lose the perceived pragmatic advantage that the lie brought in the first place. Moreover, one has also to ask the question of whether one can rely on someone telling the truth about having lied. One cannot trust them to tell the truth in this particular instance. After all, they are guilty of having lied to a group of people to gain an advantage. How is one to know that this is simply not another advantage being gained?

Very often, the lies gather by researchers are not in large enough quantities and of sufficient quality to reach any reliable conclusions about consistencies in their construction. One needs a reliable means of identifying a lie that can be demonstrated consistently over a large body and a more extended period time for conclusions to go beyond guesswork and intuition. Neither can one depend on the slow process of identifying lies within a larger to be able to establish a consistent pattern of lying and truth-telling to reach valid conclusions. In any case, this also is fraught with problems of how one might weight half-truths or things that just cannot be judged truthful or not. It is for this reason that researchers often rely on methods applied in psychological investigations where one might test a large number of individuals over a relatively short text or utterance and not a single individual over a relatively long text sample.

What is not being discussed here are lies that take place within an intimate interpersonal context, but rather, deceptive texts or utterances that take place at a relatively high institutional level. Such deceptions target an audience that does not necessarily possess the resources nor faculties to independently verify the information received nor the institution's motives for promulgating these lies. This is especially problematic in our current post-truth political reality. The public being lied to must rely on source-checking of the original suspect reports by individuals or organisations whose own honesty and journalistic veracity have been eroded by a constant barrage of statements calculated to seed doubt.

The script collection held by the National Archives in the United Kingdom consists of 267 separate daily broadcasts that start on July 3, 1943, and end on March 30, 1944. In that entire broadcast history, there are only two days that do not have a script attached to them. These 267 broadcasts consist of 207,367 words, averaging 791 words per broadcast (See figure 2). This is somewhat misleading since the early broadcasts, those before September 1943, are considerably shorter than those after that date. From September 9, 1943, onward, the average

broadcast remains relatively consistent at about 868 words per broadcast, whereas those in the two months prior have an average of 455 words per broadcast <sup>3</sup>

The date noted above, September 9, 1943, has not been chosen at random. It is with that day's broadcast that the writers of the newscasts began clearly indicating what parts of the script were true and which were fictitious.

This changes the dynamic of the document and invites the question of the motive for having marked the scripts in a way that separated fact from fiction. It was certainly not to make the job of a researcher some 80 years later easier. The PWE, and especially Sefton Delmer,<sup>4</sup> did not do such things on a whim. There must have been a reason for doing so. The reason hinted at in PWE memos is that there was a distinct possibility that they would begin to believe their propaganda. They needed a way in which they could always reference, for themselves, what was true and what was not. After all, this might have had significant repercussions up to, and including, strategic decisions made during the planning process of the eventual invasion of the continent. Moreover, they need a future reference point for their newscasts in order not to contradict their black propaganda. Though they do not articulate it, it is quite clear that they are concerned that even knowing the truth, they might believe a lie if convincingly told.

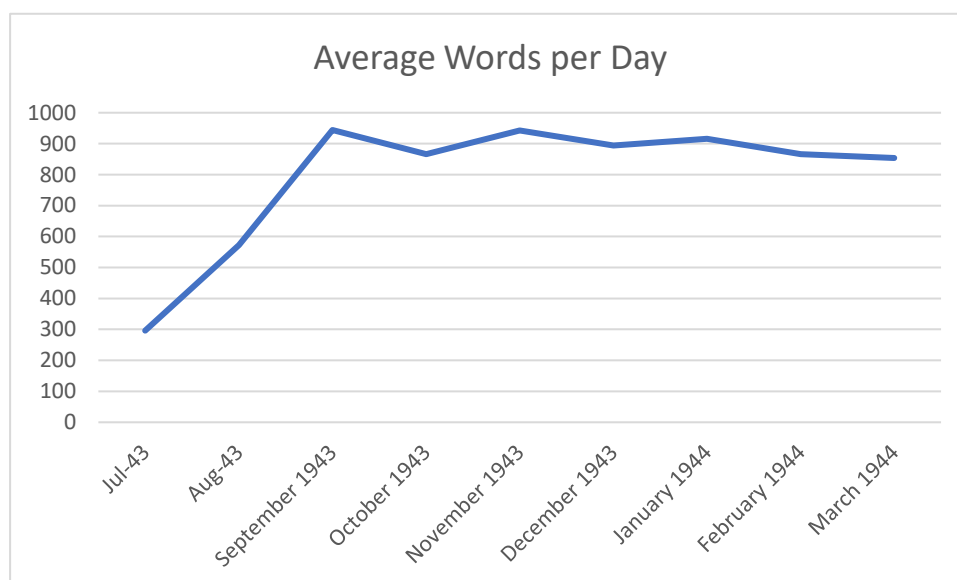


Figure 2

When concentrating on the broadcast from September 9, 1943, to March 22, 1944, some significant trends reveal themselves. For example, with one notable exception (20.12.1943), never was there more than 50% of the broadcast that consists of true statements. The upper limit of nonfiction in a broadcast was approximately 40%, with the average of true statements in a broadcast being 21.6, and with the vast number of broadcasts consisting of between 15% and 30% nonfiction with the balance being made up of known characters and fictitious events. It seems that one need only provide an audience with about 20% of verifiable information true

<sup>3</sup> Since this initial study, further broadcast scripts ending in April 1945 have been found in the National Archives at Kew (UK) that come to an additional ca. 200,000 words. These will be prepared for analysis in a subsequent research program.

<sup>4</sup> Sefton Delmer oversaw G9 Deutscher Kurzwellensender Atlantik. He was fluent in German, having grown up there, and had even been one of the foreign journalists that covered Hitler's election campaign and had even travelled with Hitler's entourage. He also led the British black propaganda efforts during the Second World War.

information to make oneself believable the rest simply needs to be a good story. However, as one reads through the script, one also realises that the concocted stories share certain qualities.

Another option that is often used in these texts is to cover up the lie within a true statement. In such a case, one takes an event that is easy to verify and then changes it with a lie that is unverifiable, but which may accentuate the natural differences that might exist between classes of people. A remarkably transparent case for this can be found in the broadcast of December 20, 1943. One hears/reads the following (the truth is italicised):

*All Wehrmacht personnel in Greece are to receive 200 cigarettes from the Reemtsma Cigarette Company's warehouses in Salonika for Christmas. SS-Oberführer Phillipp Reemtsma made a similar large gift, as a thank-offering in 1936 when the case against him for tobacco smuggling on a gigantic scale was dismissed.*

While the Reemtsma company had legal issues during the Weimar Era, the propagandists had to move the court case solidly into the Nazi-Era to make the propaganda. Moreover, the Reemtsma company, though it helped its Jewish pre-1933 partners and employees emigrate from Germany, was a consistent donor to the National Socialists.

Or,

*Professor Hunke was dismissed from all his Party offices because, in an article in the "Deutsche Volkswirtschaft," he said that the closing-down of small enterprises had only served to make a lot of people unemployed and had not improved the war effort.*

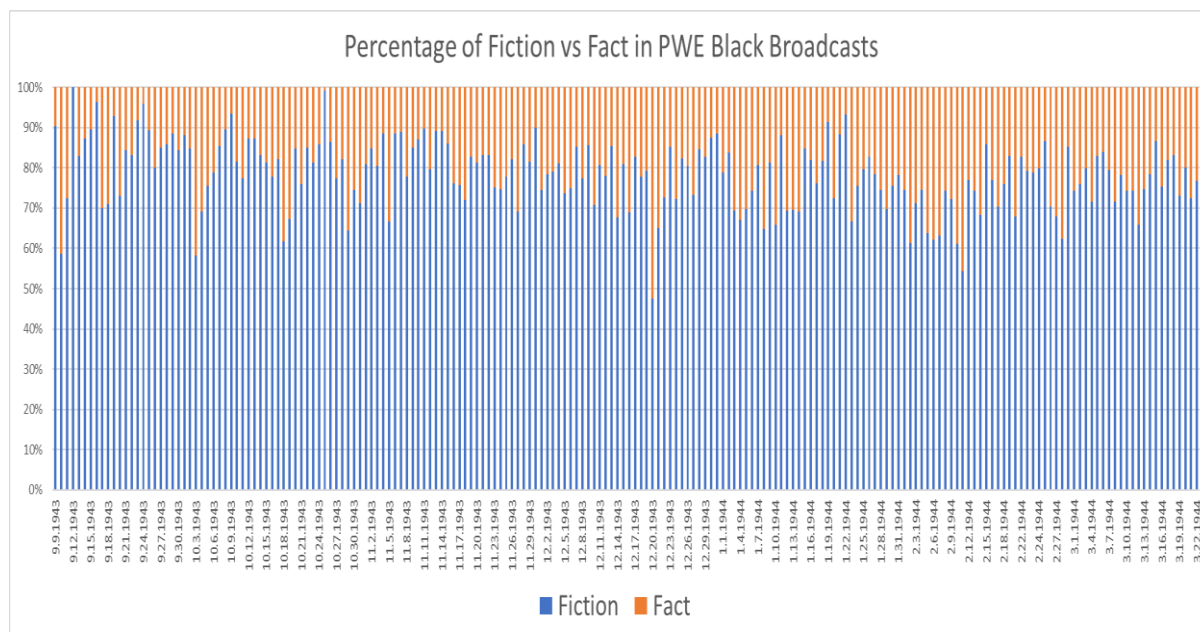


Figure 3

A good tell for a lie is if a speaker or writer provides additional pieces of information that turns a neutral statement into a negative evaluation of an individual. The same can also be done with adjectives or adverbs that might colour a statement.

While these findings are intriguing, of far greater efficacy in application to modern-day online political conversations, is the analysis of large bodies of text and the ability of sophisticated algorithms to identify personalities in text and, perhaps more importantly, subtle shifts in personalities when lies are being told. It would be helpful to determine whether it is

possible to see variations in people's personalities as they engage in telling a lie online. Up until now, researchers have had to rely on small sample groups and cobbling together various findings that they may have gathered from university studies. However, there is another source that can be applied to the material available in the London archives.

IBM, with its artificial intelligence engine Watson, has been exploring how it can sort through big data to make determinations about personality and tone in online conversations. Their primary interest is in determining who they might be able to sell something to, or, at the very least, what their customers might be able to sell to these individuals. It is also used in the analysis of customer-service interactions. In other words, they are looking for a means of targeting advertising and keeping existing clients happy.

This tool, however, has another use. It essentially administers a series of personality tests on a text, any text. The larger the text, the more accurate its findings. IBM has spent considerable resource in ensuring that their platform is reasonably accurate and has made sure that the science involved is solid. This research was begun in the early 1990s when IBM, Microsoft, Apple and other tech start-ups, began to hire applied linguists to help them with their algorithms and to hopefully make their applications, software, and platforms more user friendly. Essentially, they wanted to ensure that an ordinary individual might be able to operate their computer using something that approaches natural language.

An offshoot of all of this was the development of a processing engine that could deal with larger amounts of text and then determine the personality traits of the writers of those texts. This is, in essence, what has been done in this project. The project took texts produced by the PWE and "Soldatensender Calais" and allowed IBM's Watson to analyse the texts. The variable that was introduced was the separations of known true statements from the PWE identified false statements. This allowed the algorithm to determine the personalities of both and see if there are consistencies and inconsistencies in the personality of the text. The phrase "personality of the text" is employed because the creators of these texts were professional journalists and writers as well as academics who quite purposely and with forethought created lies that were intended to mislead large groups of people in precise ways. In addition, they purposely worked towards the production of consistent texts with which the listener could feel familiar. There was one chief editor, as well as four additional writers, and they remained constant throughout the history of the broadcaster; thus, the "personality of the text" should remain relatively stable as well.

What this project was looking for was whether there were any differences in personality between the text that represents fictional statements and those which represent factual statements. With a total of over 200,000 words over 169 separate broadcasts, one should have a reasonably consistent result if there is no difference. A further consideration is that an individual demonstrates a different type of personality in various kinds of writing. For example, when someone is writing fiction, their character will differ from when they are merely writing everyday prosaic texts to colleagues or co-workers.

The results of this test were fascinating. As demonstrated in the graph, in no small degree, personality traits do not vary. However, some telling differences indicate that certain personality traits may be necessary to tell a lie or create a believable fabrication.

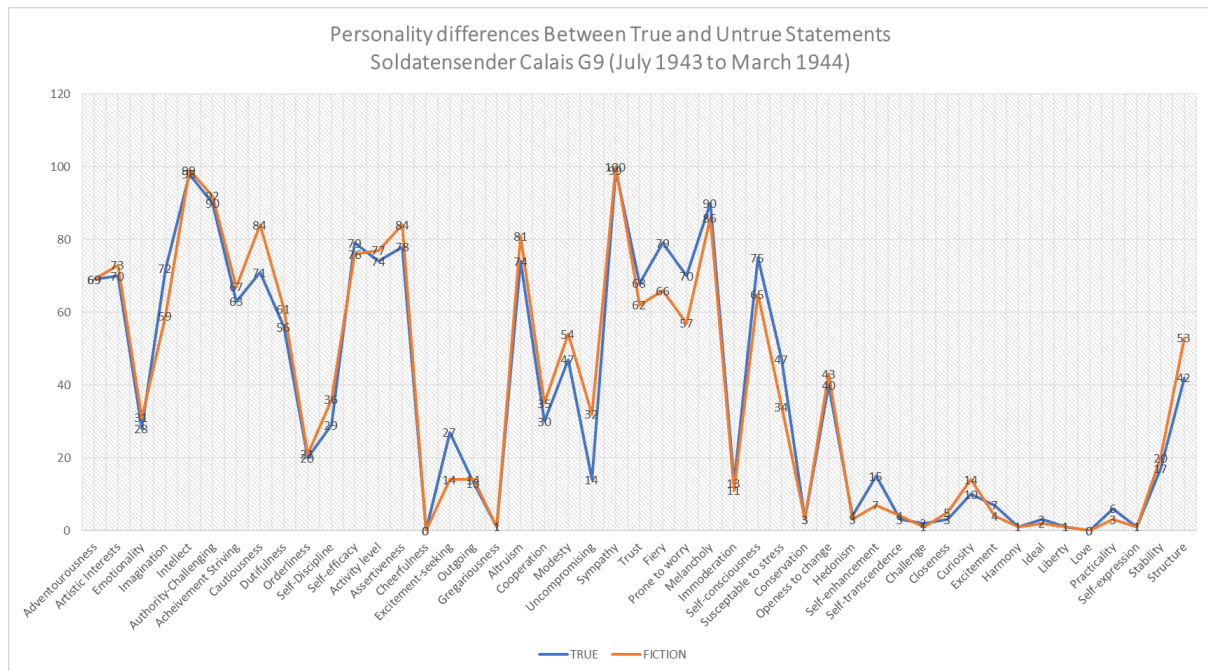


Figure 4

The Big Five personality traits theory, also known as the five-factor model (FFM) and the OCEAN model were developed in the 1980s and 1990 and are used by many big-data companies to evaluate customer interactions and by political organisations to evaluate commentary of current events and political activities.

Big Five Analysis of Personality Traits		TRUE	FICTION
		Value out of 100	Value out of 100
Openess	Adventurousness	69	69
Openess	Artistic Interests	70	73
Openess	Emotionality	28	31
<b>Openess</b>	<b>Imagination</b>	<b>72</b>	<b>59</b>
Openess	Intellect	98	99
Openess	Authority-Challenging	90	92
Conscientiousness	Achievement Striving	63	67
<b>Conscientiousness</b>	<b>Cautiousness</b>	<b>71</b>	<b>84</b>
Conscientiousness	Dutifulness	56	61
Conscientiousness	Orderliness	20	21
<b>Conscientiousness</b>	<b>Self-Discipline</b>	<b>29</b>	<b>36</b>
Conscientiousness	Self-efficacy	79	76
Extraversion	Activity level	74	77
Extraversion	Assertiveness	78	84
Extraversion	Cheerfulness	0	0
Extraversion	Excitement-seeking	27	14
Extraversion	Outgoing	13	14
Extraversion	Gregariousness	1	1



Agreeableness	Altruism	74	81
Agreeableness	Cooperation	30	35
Agreeableness	Modesty	47	54
Agreeableness	Uncompromising	14	32
Agreeableness	Sympathy	99	100
Agreeableness	Trust	68	62
Emotional Range	Fiery	79	66
Emotional Range	Prone to worry	70	57
Emotional Range	Melancholy	90	86
Emotional Range	Immoderation	13	11
Emotional Range	Self-consciousness	75	65
Emotional Range	Susceptible to stress	47	34

Figure 5

The results of a comparison between the factual and fictional parts of the broadcast reveal that the general personality between the two remains relatively consistent. This may be seen in figure 4. However, some notable traits do show consistent variation. These have been highlighted in figure 5.

When one lies,

- More cautious
- A little more self-discipline
- Cooperative Language
- One is somewhat more modest
- The language is more uncompromising

When tells the truth,

- One can be far more imaginative
- Language is fierier
- It exhibits more worry and stress
- It is more self-conscious

This preliminary study is a beginning for a better understanding of how personality "tells" on the those who create misinformation. One might even want to use the term "fake news" in this case, but that term has been so abused in the last four years that it might not be beneficial in this context. At least we have the beginnings of a method of determining the source of "fake news" and secret social media interference. In part, we owe the PWE, and their fear of losing track of telling a believable lie, for a better understanding of how we might untangle the labyrinth of misinformation that fill our social media platforms today.

---

**REFERENCES**

- Arciuli, Joanne, David Mallard, and Gina Villar. “‘Um, I can tell you’re lying’: Linguistic Markers of Deception Versus Truth-Telling in Speech.” *Applied Psycholinguistics* 31 (2010) 397-411.
- Chen, Yimin, Niall J. Conroy, and Victoria L. Rubin. “Misleading Online Content: Recognizing Clickbait as ‘False News’.” *WMDD’15* (2015), 15-19.
- Fitzpatrick, Eileen, Joan Bachenko, and Tommaso Fornaciari. *Automatic Detection of Verbal Deception*. Morgan & Claypool, 2015. DOI 10.2200/S00656ED1V01Y201507HLT029
- Girlea, Codruta, Lyft Inc and Roxana Girju. “Decoding the Perception of Sincerity in Dialogues.” *IEEE Transactions on Affective Computing* 1 (2010) 1, 1-13.
- Hauch, Valerie, Iris Blandón-Gitlin, Jaume Masip, and Siegfried L. Sporer. “Are Computers Effective Lie Detectors? A Meta-Analysis of Linguistic Cues to Deception.” *Personality and Social Psychology Review* 19 (2015) 19(4), 307–342.
- Kirner-Ludwig, Monika. “Creation, dissemination and uptake of fake-quotes in lay political discourse on Facebook and Twitter.” *Journal of Pragmatics* 157 (2020) 101-118.
- McGlynn, Joseph and Matthew McGlone. “Lying.” *Encyclopaedia of Deception*. Ed. Timothy R. Levine. Sage Publishing, 2014. Pp. 581-585.
- Meibauer, Jörg. “The Linguistics of Lying.” *Annual Review of Linguistics* 4 (2018) 357-375.
- Political Warfare Executive. “PWE/Sibs German Files: G.9 Rumours.” The National Archives (London).
- Repke, Meredith A., Lucian Gideon Conway III, and Shannon C. Houck. “The Strategic Manipulation of Linguistic Complexity: A Test of Two Models of Lying.” *Journal of Language and Social Psychology* 37 (2018) 1, 74–92.