ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis in Behavioural Economics

The Role of Knowledge and Overconfidence in a Field on the Illusory Truth Effect

Name Student: Bruno Sonnenschein

Student ID number: 597126

Supervisor: Marine Hainguerlot

Second assessor: Francesco Capozza

Date final version: 05/10/2022

Table of Contents

Abstract

In this thesis, the relationship between knowledge-based familiarity with a field and the Illusory Truth Effect is tested and discussed. The Illusory Truth Effect generally states that repeating statements increases their credibility, regardless of their actual truth status. Previous literature has been divided on the importance of knowledge on this effect. Knowledge-based familiarity may either reduce or remove the effects of repetition or could instead strengthen them, depending on the circumstances. Findings from this study indicate that the impact of familiarity with a field on the truth effect, for field-related statements, may vary between levels of familiarity, and be further influenced by overconfidence. Little familiarity within an area seems to invite stronger effects of repetition on the credibility of fake news, while higher familiarity seems to deter the effects. Overconfidence, which is found especially in individuals new to a field, seems to further increase the size of the Illusory Truth Effect. This thesis, therefore, adds to and expands the existing literature and makes suggestions to policy-makers to support them in the tough battle against the spread of fake news.

Section I – Introduction

"If you are telling a big lie, and repeat it often enough, people will start to believe it" is a quote generally attributed to Adolf Hitler's propaganda minister Joseph Goebbels (Krieghofer, 2017). Interestingly enough, there seem to be no trustworthy sources that can confirm or deny that Goebbels ever said these words, and it is an ironic circumstance that a likely false quote was attributed to one of the most important propagators of misinformation and propaganda in history (Doob, 1950). Nevertheless, the contents of this potentially false quote bear an important fact, and may they themselves hold true. The theory that the mere repetition of false information increases its credibility is nowadays generally accepted and many studies have found and discussed this so-called *Illusory-Truth-Effect* (Dechêne et al., 2010), further also referred to as the *IT-Effect* or simply *Truth Effect*.

It is an easy concept to understand that a lie if repeated often enough, can become a perceived truth in our minds. In our everyday lives, we tell ourselves and the people surrounding us a wide range of lies, big or small, hurtful or harmless, and while some of these lies may have big impacts on our or other's lives, these are arguably drops in the ocean compared to lies that affect society as a whole. The lies that are most damaging, are the ones that affect not only one person but rather a whole country or even the entire world. Throughout history and ever so in recent times, propaganda, disinformation, and fake news have been used as tools to steer populations toward certain beliefs (Brandenberger, 2012; Voigtländer and Voth, 2015). Most recently the American elections, surrounding Donald Trump, as well as the Covid-19 pandemic have seen blatant fake news running rampant in all forms of media (Allcott and Gentzkow, 2017; Patwa et al., 2021; Furini et al., 2020). Especially the internet and social media have been playing a big role in the success of fake news. With the formation of so-called "bubbles" and the polarization of beliefs strengthened by algorithms (Alcott and Gentzkow, 2017), people find themselves continuously confronted with false information with few chances to avoid it and few reasons to not believe it. Social media algorithms make sure that individuals are always surrounded by others that share their beliefs, and information that confirms them, thereby potentially turning doubts into certainty for both true and false statements, leading to a polarization of society and the appearance of more extreme beliefs (Tucker et al., 2018).

In light of this major struggle against fake news, it is important to understand how and why certain misinformation is effective. By knowing the mechanisms behind the success of false information, combatting said issues and correcting objectively false beliefs should become easier and more efficient.

The goal of this thesis and the research that underlies this work, is therefore to further understand one aspect of the effectiveness of fake news. This aspect is the aforementioned IT-Effect, in other words, the idea that merely repeating statements, true or false, will increase their credibility. Specifically, the aim is to shed light on the relation of familiarity in a field and its effects on this phenomenon, concerning statements from the same field. By further understanding the role of knowledge and familiarity with a topic and its relationship with the IT-Effect, important insights can be gained to improve the ongoing combat against fake news. For example, if it is shown that an increase in familiarity is linked with a reduction of the IT-Effect, information campaigns may prove useful tools to fight fake news. If, however, familiarity is seen to increase the size of the IT-Effect, such information campaigns may prove counter-effective and worsen the outcome.

Previous research into the Illusory Truth Effect has shown diverse results regarding both the effect of specific knowledge pertaining to a statement and simply regarding familiarity with a statement. In this case, when talking about specific knowledge, the idea is that individuals have the direct knowledge in their memory that should theoretically allow them to correctly deduct the truth status of a statement. Relatedly, some authors claim that knowing the true answer to a question will deter the IT-Effect (Dechêne et al., 2010), and others somewhat refute this assumption, showing that in some cases, even individuals who should know and do know better, can be affected by the IT-Effect (Fazio et al., 2015). Fazio et al. (2015) measured this knowledge by performing a knowledge check in which subjects had to answer questions directly rather than judging the credibility of statements on a scale. In cases in which individuals correctly answered the questions in the knowledge check, the corresponding statements were marked as “known” and pre-existing specific knowledge was assumed.

In the broader case, when talking about circumstantial knowledge, field-specific knowledge, or simply familiarity with a statement, the idea is that individuals may have knowledge or experience in a topic that could help them make a correct truth judgement regarding a statement. Consequently, it is different to specific knowledge due to the fact that the exact

knowledge regarding the truth status of a statement does not necessarily need to be present in the person's memory. Therefore, in the context of this study, such knowledge will primarily be referred to as familiarity with statements from a certain field.

In regards to such familiarity within specific fields prior studies have also had somewhat contradicting results concerning its relationship with the size of the Illusory Truth Effect. Some studies found a smaller IT-Effect for individuals with high familiarity (Srull, 1983), while others found high familiarity to increase the IT-Effect (Boehm, 1994; Arkes et al., 1989). However, it remains unclear what drives this difference in results. Variation in the findings could be the result of an inherent difference in the IT-Effect between certain fields, or a difference in how familiarity was measured and established in the related studies. Also, a too broad categorization of familiarity may not differentiate between overconfidence and actual knowledge, thus biasing the findings. Therefore, the aim of this study is not only to find more evidence of the effects of familiarity with a topic on the IT-Effect but also to understand the exact relationship between the two. Further, this study aims to find factors, such as differences in measuring knowledge and familiarity that may potentially influence the results.

Concretely, the research question that I will attempt to answer is, "Does familiarity within a field increase the size of the Illusory Truth Effect for statements from said field, and is overconfidence in the said field a factor that can explain this phenomenon?".

This study will therefore add to the existing literature by attempting to clarify the relationship between familiarity with a topic and the IT-Effect, and examine the differences between prior studies. Additionally, it will try to deliver information on the role of overconfidence in said relationship.

To answer the abovementioned research question, I perform an online experiment. The experiment is similarly structured to previous experiments looking at the Illusory Truth Effect, and in addition, includes measures that give information on subjects' familiarity within certain fields. Concretely, the experiment includes an exposure phase in which subjects are shown statements for the first time, followed by a knowledge check in which the subjects' levels of familiarity in the fields of economics and medicine are estimated via a short quiz. Finally, subjects are again exposed to a series of statements, half of which are new to them and the other half are drawn from the first series of statements. In this second phase of exposure,

subjects are then asked to judge the truth status of each statement. The included statements are related to the field of economics, and medicine, and a few are unspecified. Via this procedure, I am able to calculate and test for differences between the average truth ratings of repeated and new statements overall and for both specified sectors.

Subsequently, using the information on the truth ratings, coupled with the estimated levels of familiarity within the fields of interest, I am able to conduct tests that look at potential systematic differences in the size of the IT-Effect for subjects with high and low familiarity. Furthermore, by asking subjects to rate their confidence when answering the questions included in the quiz, I am able to judge their degree of overconfidence in the relevant field and consequently test for differences in the IT-Effect between individuals with a tendency to be more overconfident and those who are less overconfident.

Overall, the main body of this thesis is structured in the following way. In section II, the theoretical foundation of the IT-Effect and the mechanisms surrounding it will be discussed. Current findings will be presented and juxtaposed, and the basis for my research question and hypotheses established. Section III will present the details surrounding the experimental study and thus the data used for the analysis. In section IV results will be presented and hypotheses tested. Section V will discuss the theoretical analysis based on the findings. Finally, section VI will summarize the findings and discuss the limitations of both the performed study as well as the drawn conclusions. Further, there will be a general discussion of the results, their implications, and their usefulness in potentially combatting fake news. Suggestions for mechanisms to fight misinformation will be provided and possible goals for future research will be discussed.

Section II – Theoretical Background

In this section I will discuss the relevant theoretical background on which the analysis and argumentation are built. I will discuss the relevance of the research, the potential functioning of the Illusory Truth Effect, and the related potential role of familiarity and overconfidence.

First, I establish the status quo and the primary sources of fake news in the modern age. Clarifying the setting in which fake news are encountered the most, allows me to make better suggestions for combatting misinformation and the IT-Effect.

In a second step the core mechanism behind the Illusory Truth Effect, namely fluency, is explained and discussed. By looking at the findings of prior studies into the IT-Effect, a theoretical foundation is established which will help with further discussion and analysis.

Following, the role of knowledge and familiarity in regards to the IT-Effect is reviewed based on previous research. The opposing findings regarding this relationship are discussed, and an attempt is made to reason how these results came about and what can explain their seemingly contradicting nature. A point is made that a variation in results may be the consequence of different methods used to classify the levels of familiarity in a field.

Fourth, in direct relation to the previous step, the potential role of overconfidence in influencing the relationship between familiarity and the IT-Effect is considered. An argument is made that overconfidence among beginners may be the main driver of the Illusory Truth Effect found, by previous studies, in individuals classed as familiar with a topic.

In a fifth section, the currently and formerly employed and tested measures against fake news are presented. By understanding which measures seem to work and which do not seem to work, I again clarify the importance of the current research. Furthermore, I establish potential negative consequences that could come from policies that do not sufficiently take into account the findings regarding the IT-Effect.

Finally, I discuss the concrete hypotheses that will be tested, and that should allow me to draw concrete conclusions concerning the relationship between familiarity in a field and the IT-Effect.

II.1 – Fake News and Social Media

The existence of fake news, or more generally the spread of misinformation is by no means a recent phenomenon. In past centuries strategies of indoctrination and propaganda were used to steer public opinion and control beliefs. Such measures were easily implemented in countries such as the USSR under Stalin or Germany under Adolf Hitler (Brandenberger, 2012; Voigtländer and Voth, 2015) in which disinformation strategies were led by the governments themselves. However, in more recent years, in countries with high regard for democracy and freedom of speech, indoctrination or other methods of changing beliefs on a similar scale are hardly possible anymore. Instead, individuals or groups with certain objectives may employ less intrusive measures to influence beliefs, for example, fake news. Freedom of expression and freedom of the press not only allow for the truth to be spoken without risking persecution but also allow for falsities to be spread with few tools to deter them.

The likely largest contributors to the spread of fake news in our modern society are social media platforms. In 2016, 62% of US-American adults got their news from social media platforms (Gottfried and Shearer, 2016). In the same year, Donald Trump won the presidential elections amongst worries about fake news. In their study, Alcott and Gentzkow (2017) reviewed the relevance of fake news in the US presidential elections. They found that misinformation was spread about both presidential candidates but pro-trump fake news was quantitatively more prominent. Further, they established that each US adult may have encountered at least one or two fake news stories during the elections, suggesting that exposure to misinformation was widespread. In a more up to date report (Shearer and Mitchell, 2021), the number of US-American adults who got their news from social media at least occasionally is said to have risen to 71%.

In more recent years, the Covid-19 pandemic saw another uprise in fake news (Moscadelli et al., 2020; Fernández-Torres et al., 2021; Wang et al., 2019). False information about the side effects of vaccines, the effectiveness of mask usage, and potential alternative treatment methods were widely spread. Consequently, while fake news would usually only affect the ones that believe them, or at worst lead to political turmoil, they would not pose an imminent physical danger to the general population. However, with fake news entering the realm of medicine, misinformation poses a danger not only to its believers but also to those around them. For example, the scepticism surrounding vaccines would not be an issue in a scenario

in which everyone could be vaccinated and the vaccines were fully effective. In such a case only those who chose not to vaccinate themselves would be in danger. However, in reality, the success of vaccines relies on a certain share of the population participating. Herd immunity is necessary to protect those that are not able to be vaccinated due to various possible reasons (Randolph and Barreiro, 2020). Consequently, the spread of fake news in the medical sector poses an important societal danger, and in regards to vaccines, it puts especially all those at risk that do not have a choice.

It is thus important to further understand how fake news spreads, and what factors make it spread so efficiently. Moravec et al. (2018) found that most users of social media may be better at detecting fake news by throwing a coin to guess the truth status rather than trusting their beliefs. This general difficulty for people to correctly judge the validity of information encountered online, paired with the deliberate use of influencing techniques by propagators of fake news, creates an environment in which misinformation is rampant and effective.

II.2 – Fluency and the Illusory Truth Effect

One of the potential factors that can influence the success of fake news is repetition. This so-called Illusory Truth Effect, first found by Hasher et al. (1977), coined by Begg et al. (1992), and also referred to as the validity effect (Boehm, 1994) or simply the truth effect (Dechêne et al., 2010), is a well-known phenomenon in psychology. It has been found to appear in various circumstances and environments, and for statements from a wide range of areas (Hasher et al., 1977; Bacon, 1979; Dechêne et al., 2010). In short, the theory is that mere repetition of a statement increases its credibility, regardless of whether it is true or false, and regardless of whether it is implausible or plausible (Fazio et al., 2019).

It is also important to note that in general, the IT-Effect seems to persist across both short (Begg et al., 1985) and long (Bacon, 1979) periods of time. It is thus a valid concern that repetition of fake news may lead to lasting false beliefs among the general population.

To better understand this effect, various researchers have taken different approaches and evaluated possible explanations for this mechanism. One major point, deemed to be the main explanatory aspect for the IT-Effect regardless of other potential factors, is the role of processing fluency (Dechêne et al., 2010). Processing fluency describes the ease with which individuals understand and process a certain piece of information. In situations in which

persons do not rely on or lack the necessary objective knowledge to accurately judge the truth status of a statement, they tend to rely on fluency (Alter and Oppenheimer, 2009).

Factors that may improve fluency include, but are not limited to, language, length, likeability, visual representation, and familiarity (Alter and Oppenheimer, 2009). For example, Reber and Schwarz (1999) found that statements presented in an easy-to-read font were rated as more likely to be true compared with statements that were difficult to read. It follows therefore, that the repetition of statements, an action that improves the familiarity of ideas and notions and that may also help with understanding difficult language, should improve the processing fluency (Unkelbach, 2007; Lewandowsky et al., 2012). In other words, if we see a statement often, it becomes easier to understand. And more fluent statements, are also more likely to be believed and rated as true (Begg et al., 1992; Schwarz, 2004). Relatedly, evidence from neurology has shown links between the brain regions associated with processing fluency and the Illusory Truth Effect (Wang et al., 2016), thus further emphasizing the central role of fluency.

Furthermore, De Keersmaecker et al. (2020) investigated whether individual differences in cognitive ability, need for cognitive closure, and cognitive style affected the mechanism behind the IT-Effect. According to the authors, all three factors are likely to influence the role of processing fluency. Cognitive ability, or intelligence, is said to determine memory processes which in turn impact information processing. A high need for cognitive closure, or the desire for firm answers, is said to be associated with higher reliance on intuition and heuristics. And the cognitive style, or the fact of thinking more intuitively or more analytically, can influence the reliance on fluency. However, De Keersmaecker et al. (2020) found no conclusive evidence of a relationship between the three mentioned factors and the Illusory Truth Effect, thus suggesting that perhaps some other individual differences may be more relevant in determining the size of the effect.

In regards to familiarity, it is important to distinguish between different types. Familiarity in general means that the presented information seems known to a subject, and evokes some type of either true or artificial recollection. Familiarity can therefore stem from prior knowledge, as will be explained in the following section, or it can stem from actual or even make-believe previous exposure (Bacon, 1979). Thus, when discussing the effects of

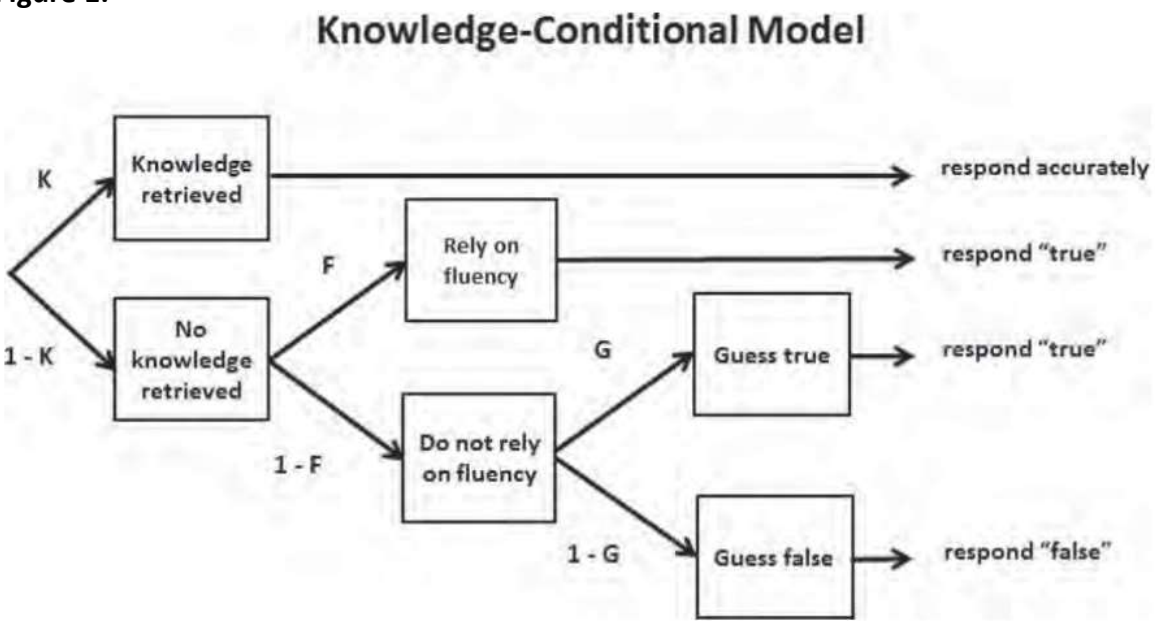
familiarity in the scope of this thesis, I will strictly be referring to familiarity due to actual or perceived pre-existing knowledge prior to the experiment.

II.3 – Knowledge and the Illusory Truth Effect

While the role of processing fluency in the effectiveness of the IT-Effect is generally agreed on (Unkelbach, 2007), the role of knowledge is somewhat more controversial. While some papers agree that knowing the true state of some piece of information should remove the IT-Effect (Dechêne et al., 2010) and that reliance on knowledge precedes that on fluency (Fig. 1), other studies find that the IT-Effect, and more specifically fluency, may in some cases overpower pre-existing knowledge and to some degree override previous beliefs temporarily or entirely (Boehm, 1994; Fazio et al., 2015).

Fazio et al. (2015), suggest that people do not generally process statements using their inherent knowledge first, but instead may act more instinctively by relying on the fluency of a statement to rate its credibility. This so-called fluency-conditional model (Fig. 2) is therefore in direct contradiction to the so-called knowledge-conditional model (Fig. 1). The figures show that there is a certain order in which information is supposedly processed, either based on fluency first or based on knowledge first.

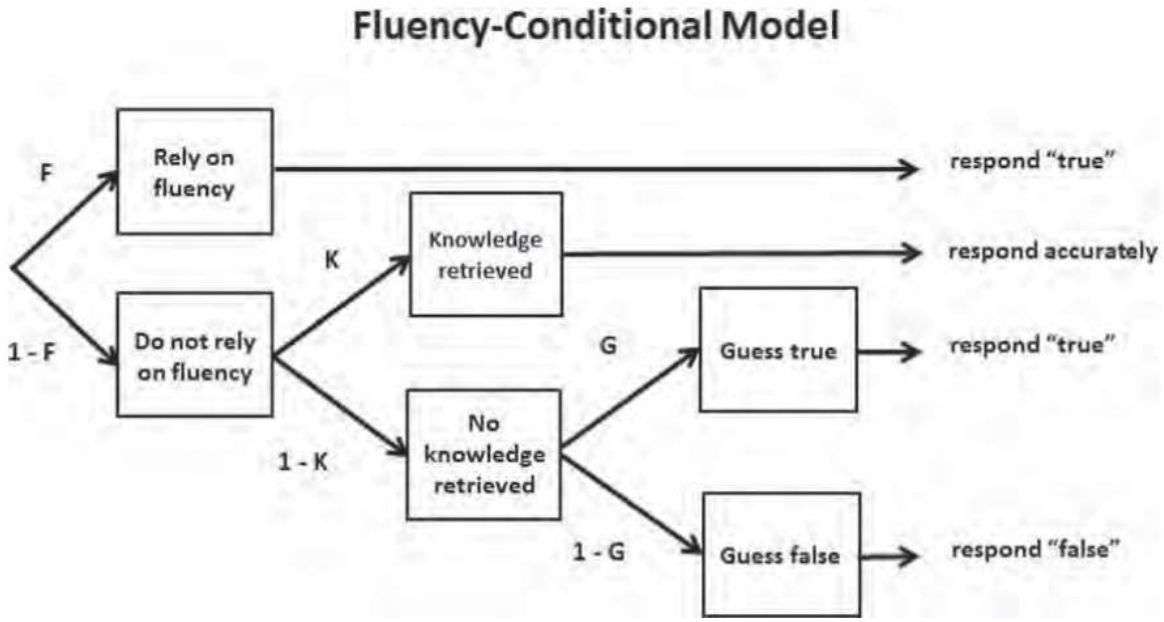
Figure 1:



Note: The Knowledge-Conditional Model by Fazio et al. 2015. The model portrays the general mechanism behind the Illusory Truth effect as described in previous research such as Dechêne et al. (2010), with knowledge having a more concrete assumed role in deterring the effects of repetition on validity judgments. Each Branch has a probability (0-1) to take effect, with K being the probability to rely on knowledge, F the probability to rely on fluency, and G the probability to "Guess True". Thus, with probability K, a person will accurately rate the truth

status of a statement, with probability 1-K * F a person will rely on fluency and thus rate the statement as true. With probability 1-K * 1-F * G and 1-K * 1-F * 1-G the statements will be guessed to be true or false respectively. From "Knowledge Does Not Protect Against Illusory Truth" by Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J., 2015, Journal of Experimental Psychology: General, 144(5), p.997. Copyright 2015 by the American Psychological Association.

Figure 2:



Note: The Fluency-Conditional Model by Fazio et al. (2015, p.997). The model portrays the updated mechanism behind the Illusory Truth Effect, as described by Fazio et al. (2015), with fluency being the primary factor for judgments on validity. Each Branch has a probability (0-1) to take effect, with F being the probability to rely on fluency, K the probability to rely on knowledge, and G the probability to "Guess True". Thus, with probability F, a person will rely on fluency and rate the statement as true, with probability 1-F * K a person will rely on knowledge and thus rate the statement according to the actual truth status. With probability 1-F * 1-K * G and 1-F * 1-K * 1-G the statements will be guessed to be true or false respectively. From "Knowledge Does Not Protect Against Illusory Truth" by Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J., 2015, Journal of Experimental Psychology: General, 144(5), p.997. Copyright 2015 by the American Psychological Association.

The majority of studies concerning the IT-Effect prior to the study by Fazio et al. (2015) assume the role of knowledge to be obvious (Dechêne et al., 2010; Hasher et al., 1977). The assumption made across these studies is that knowing the relevant information to properly judge the credibility of a statement will in any case negate or at least reduce the effects of repetition. Relatedly, Dechêne et al. (2010, p.239) state that "Statements have to be ambiguous, that is, participants have to be uncertain about their truth status because otherwise the statements' truthfulness will be judged on the basis of their knowledge". This means, therefore, that previous studies assumed that the IT-Effect could not be found for statements whose truth status should be known to subjects. However, the study by Fazio et al. (2015) challenged this standard assumption and surprisingly found that prior knowledge may in fact not matter as much as previously assumed.

This is somewhat in line with the theory developed by Daniel Kahneman, who distinguishes two types of thinking, system 1 and system 2 (Kahneman, 2011). System 1 thinking relies on automatic processes and is fast and unconscious, while system 2 thinking relies on an effortful, slow, and conscious process. Thus, in the case of fake news and the role of knowledge, it can be argued that individuals will rely on their system 1 thinking when briefly evaluating a statement's truthfulness, thereby making them vulnerable to subconscious effects and an improved processing fluency via repetition. This theory, and line of thinking, emphasizes the danger of the IT-Effect when, likely rightfully so, assuming that people rely on their system 1 thinking while consuming news on social media (Moravec et al., 2018).

Consequently, the theory established by Fazio et al. (2015) as well as their findings in regards to specific pre-existing knowledge, shows that the IT-Effect can even overpower prior beliefs. In fact, they found that regardless of false statements being known or unknown, the IT-Effect prevailed. Even more surprisingly in one of the two experiments, in regards to falsehoods, the IT-Effect appeared to be larger for known false statements than for unknown ones.

Referring to this theory of a fluency-conditional model (Fazio et al., 2015), it is therefore interesting to examine whether being knowledgeable in a field and thus being familiar with statements from said field, will combat the IT-Effect concerning statements whose exact truth state should be unknown even to experts in said area. If the IT-Effect can be reduced or eliminated by being knowledgeable in a field, then increasing general education in certain areas may be sufficient to combat some aspects of the spreading of fake news. If, however, even experts are prone to the IT-Effect, due to increased familiarity strengthening the reliance on fluency, then some other measures that, for example, concentrate on simply reducing the exposure to fake news may be more effective.

Previous researchers have found conflicting results when testing the effects of knowledge in specific domains on the IT-Effect of statements belonging to the same area. Srull (1983) found that individuals more familiar with a topic displayed a smaller IT-Effect than non-familiar subjects. Familiarity in this case can be seen as somewhat synonymous with the notion of field-specific knowledge. However, Arkes et al. (1989) found that the IT-Effect was not present for statements regarding topics that the subjects were unfamiliar with, and instead, only areas of high familiarity displayed an increased truth rating due to repetition. A first distinction between the study by Srull (1983) and that by Arkes et al. (1989) lies in the measurement of

familiarity. While the first study measured high and low familiarity directly via an objective test, subjects in the second study self-assessed their familiarity by ranking fields in order of most to least familiar. Boehm (1994) uses the field of study as an indicator for familiarity with a topic, and finds, similar to Arkes et al. (1989), that high familiarity with a topic seems to invite a stronger IT-Effect.

A possible explanation for this counterintuitive effect of knowledge on the impact of repetition on truth judgments could be a false familiarity. In other words, individuals that are familiar with a topic but not with the specific statements may subconsciously rely on a false sense of familiarity. Consequently, instead of depending on their knowledge to infer the correct truth state of a statement, they might intuitively establish the prior seen statement as factual. Boehm (1994, p.288) states "One can imagine beliefs being formed where the perceived message is not of sufficient interest to activate counterarguing", thus setting a condition that as long as an individual does not recollect the specific knowledge that may counter the information from the seen statement, familiarity with the statement will potentially form false beliefs.

The question that remains, however, is where exactly the conflicting results between the presented studies stem from. One possible explanation may be the varying measurements of familiarity and knowledgeability. Srull (1983) performs an objective test to classify subjects into high or low familiarity, meanwhile, Arkes et al. (1989) have subjects self-assess their level of familiarity, and Boehm (1994) uses the subjects' field of study as a proxy for being highly familiar with the respective topic.

It also remains unclear which levels of knowledge or familiarity may lead to a reduction or increase in the IT-Effect. It seems reasonable to assume that little knowledge, which simply increases familiarity, may strengthen the effects of repetition due to the improved fluency of statements. Meanwhile, high knowledge may see a less prominent reliance on fluency when encountering false information. Due to a possibly higher probability of recognizing inconsistencies in false statements, the Truth Effect for falsehoods could diminish.

Finally, due to the similarity of the research question and the underlying assumptions between this study and the paper by Fazio et.al (2015) all future arguments in this text will be based on the fluency-conditional model (Fig.2). However, it should be noted that the existence and

effectiveness of the knowledge-conditional model (Fig. 1) is not disregarded entirely. Moreso, I believe the context in which individuals generally encounter fake information the most, namely social media, to be one in which they would generally rely on intuitive thinking first and foremost (Moravec et al., 2018). In other circumstances, people may very well be more thoughtful of their beliefs and rely on their knowledge first (Fazio et al., 2015). In fact, when subjects are actively encouraged to rely on their knowledge and recollection, any potential influences of repetition seem to disappear (Begg, 1992).

Which system of thinking subjects would rely on in an online experiment is difficult to ascertain. However, I believe that while subjects are generally expected to do their best, they will also attempt to complete the experiment in the shortest time possible, thus making it more likely that they would primarily rely on fluency rather than knowledge. Consequently, the experimental environment should somewhat replicate the usual real-life environment in which fake news is encountered, and in which the IT-Effect poses a danger.

II.4 – Overconfidence and the Illusory Truth Effect

In the scope of this study, a series of factors that may influence the differing findings as to the impact of knowledge on the IT-Effect will be examined. One such potential factor is overconfidence. In the context of this thesis, overconfidence refers to the first definition by Moore and Healy (2008), namely so-called overestimation. This definition covers situations in which subjects overestimate their "actual ability, performance, level of control, or chance of success" (Moore and Healy, 2008, p.3).

It seems that one potential reason which explains the increase in the size of the Truth Effect for individuals who are supposedly familiar with the statements could lie in the relationship between knowledge and overconfidence. Sanchez and Dunning (2018) explored how overconfidence developed in subjects with increasing levels of knowledge and found that especially individuals who are new to an area are likely to be overconfident. Consequently, when assuming that overconfident people are more likely to be affected by repetition effects, the differences in results between the studies by Srull (1983) and Arkes et al. (1989) may lie in the variance in knowledge levels and thus also in the differing levels of overconfidence between individuals.

In the study by Srull (1983) subjects who were classified as highly familiar with the topic, were placed there via relatively objective measures and could be seen as truly knowledgeable in the relevant field. However, in the study by Arkes et al. (1989), subjects' knowledge levels were classified via highly subjective means. Thus, if one were to conclude that the first study measured the IT-Effect for experts, and the second study measured the IT-Effect across a more varied and likely less knowledgeable field of subjects, then the differences in their findings could be partially explained by the underlying variance in overconfidence in individuals of varying expertise. Consequently, the high familiarity individuals in the study by Arkes et al. (1989) may consist of a mix of truly familiar subjects as well as subjects that are overconfident and merely believe themselves to be familiar with the topic.

The direct link between the IT-Effect and overconfidence has not been at the centre of attention for past and current researchers, and thus any direct evidence or supported claims are scarce. However, concerning fake news in general, it has been found that more overconfident individuals will tend to be more susceptible to fake news, and be quicker to believe unsourced statements (Lyons et al., 2021; Pennycook and Rand, 2020). The link between the increased susceptibility and overconfidence can potentially be attributed to, as Pennycook and Rand (2020, p.186) put it “reflexive open-mindedness”, meaning that they are generally quicker to accept facts without reflectively thinking about them.

When applying this idea about "reflexive open-mindedness" to the established fluency-conditional model (Fig.2), it seems intuitive to believe that overconfidence will lead individuals to rely even more on fluency rather than knowledge when making swift judgments. Additionally, prior exposure may establish statements as factual in the subconscious minds of overconfident subjects. Thus, when being exposed to statements a second time, the previous information will be referred to as the next-best information to base the truth status of the statement on. It would therefore follow, that overconfidence in individuals should increase the effects of repetition on their judgment of truth.

Adding to this, Lyons et al. (2021) find some concerning links between overconfidence and being not only worse at differentiating false from true information but also being more likely to share false information and refer to untrustworthy sources. This is concerning as it shows that people who are more susceptible to fake news, may also be the ones to be confronted with it the most. A likely stronger IT-Effect, paired with the aforementioned role of social

media algorithms and bubbles, may thus create a dangerous environment for the effective spread of misinformation.

II.5 – Potential Measures against Fake News

The underlying reason behind this study and likely all other research concerning the IT-Effect is to further understand the driving forces of fake news in order to be able to combat them. While understanding human behaviour is interesting by itself, I believe that the driving factor of all research is to improve our society. Therefore, it is important to discuss all current measures against fake news, as well as which measures may potentially arise or prevail based on the research on the IT-Effect.

Various methods to prevent the spreading of fake news have been tested and implemented in the past years. If one frequents a number of social media platforms, one has likely seen informational banners attached to certain posts. For example, in the context of the COVID-19 pandemic, on various platforms, all posts that mentioned the virus or pandemic in some way or another were accompanied by both a brief informational message concerning potential misinformation, as well as a direct link to a governmental or otherwise official website containing clarifying and sourced information. In cases in which a post contains proven false information, platforms may even block the content entirely or cover it with a fake news warning that must be circumvented to see the post in question (Mosseri, 2016). However, it is almost impossible to fully accurately attach fake news warnings to all posts where it matters and where it is appropriate. In other words, not all misinformation is labelled as such, and some true information may wrongfully get warnings.

Several studies explored the effectiveness of different measures to reduce the spread of fake news. For one, affirming facts appears more effective than retracting myths in regard to correcting long-term beliefs (Swire et al., 2017). Additionally, including the concrete piece of misinformation in a retraction seems to be more effective than less explicit retractions (Ecker et al., 2017). Flagging headlines as false seems to only have a limited influence on user's beliefs (Moravec et al., 2018; Pennycook et al., 2018) and appears to increase the perceived validity of other non-flagged news, regardless of their truth status (Pennycook et al., 2020). The effectiveness of warnings of potential fake information seems to be weakened by pre-existing familiarity, thus suggesting that early warnings and clarifications of fake information are crucial to stopping the rise of fake news (Pan and Hu, 2022). However, it was also found that

informing individuals of fake news after the first exposure is more effective than doing so prior to or during the first exposure (Brashier et al., 2021). Furthermore, when debunking misinformation, the level of detail and completeness of the debunking message is important. Corrections that only give slight encouragement to consider the counterpoint appear to strengthen initial misconceptions, while more in-depth corrections of the false information successfully reduce the belief in fake news (Chan et al., 2017).

Additionally, several studies warn of a repetition effect caused by warning and debunking (Swire et al., 2017; Pan and Hu, 2022; Lewandowsky et al., 2012). Due to the functioning of the IT-Effect and its reliance on familiarity rather than concrete memory, it seems that even prior exposure to warnings of certain misconceptions may in the long run evoke an IT-Effect by making individuals more familiar with the false information (Pan and Hu, 2022).

However, this so-called continued-influence effect is also controversial, and several studies have found some evidence for the contrary (Ecker et al., 2017; Wahlheim et al., 2020), although they did not test for the importance of delays and the persistence of effects over time. Skurnik et al. (2005) found that in their study, repeatedly classifying statements as false improved the accuracy of truth judgments in the short term, but led to increased truth ratings for the same false statements after a three-day delay. This effect was especially prominent in older subjects, who are evidently already more prone to share and believe fake news (Brashier and Schacter, 2020).

Consequently, it still remains somewhat unclear whether repeating misinformation is more beneficial or hurtful when correcting beliefs in both the short and the long term. Further understanding of the IT-Effect and its various influencing factors, should therefore also help with the diverse evidence concerning the correction of false beliefs.

II.6 – Hypotheses and Theoretical Approach to Research Question

To answer the underlying research question, and provide evidence that confirms or denies the proposed series of argumentation, a range of hypotheses will be tested. The hypotheses cover the Illusory Truth Effect regardless of differing degrees of familiarity, the relationship between familiarity in a field and the size of the IT-Effect, as well as the potential influence of overconfidence on the IT-Effect.

Given the previous research and the large evidence towards the IT-Effect, I expect my data to also confirm a general effect of repetition on the perceived credibility of both true and false statements. All hypotheses assume that all the included statements fulfil the conditions of being similarly fluent and similarly believable. Thus follow my first two hypotheses.

Hypothesis 1: "On average, repeated true statements receive higher truth ratings than new true statements"

Hypothesis 2: "On average, repeated false statements receive higher truth ratings than new false statements"

Going from the more general hypotheses to the hypotheses regarding knowledge, the third and fourth hypotheses are based on the assumption that a higher general knowledge in a specific field, and thus a high familiarity with statements from said area, will make a person less prone to the IT-Effect for false statements from the said field. I expect it to be the case that high familiarity with the relevant topic activates the more conscious system 2 thinking when confronting individuals with statements that are false, regardless of the previous repetition.

Consequently, I believe that repeating false statements from a field with which a person is highly familiar, is more likely to make the individual rely on a more conscious belief system. This is because, in a first exposure to the false information, the individual may already consciously or subconsciously realize some contradiction in the presented statement. Therefore, when the statement is repeated that doubt will actualize itself and lead to a more knowledge-based approach to the truth rating. I thus expect familiar subjects to opt for a system 2 thought system when confronted with false statements. This can be visualized using the aforementioned "Fluency-Conditional Model" schematic by Fazio et al. (2015, p.997; Fig.2). The idea is that the contradiction, made apparent due to the high familiarity, will trigger a more conscious process in which doubt will lead the subject to not rely on fluency, and thus act on their inherent knowledge as much as possible. The exact hypothesis is therefore as follows.

Hypothesis 3: "For field-specific false repeated statements, individuals being highly familiar in the relevant field, will on average see a significant reduction of the Illusory Truth Effect compared to individuals that are less familiar with the topic."

While hypothesis 3 covers false statements in the presence of high familiarity in a field, I believe the effects on true statements to be different. Since I am arguing that false information triggers system 2 thinking, it follows that such a triggering should not occur when subjects are exposed to true statements. Since there should be no conflicting information, subjects should continue to rely on fluency alone when judging both repeated and new true statements. In this case, the additional familiarity with the statements should strengthen the effects of repetition and thus lead to a stronger IT-Effect. Therefore, hypothesis 4 is as follows.

Hypothesis 4: "For field-specific true repeated statements, individuals being highly familiar in the relevant field, will on average see an increase to the size of the Illusory Truth Effect, compared to individuals that are less familiar in said field."

To test the prior assumptions that subjects with high familiarity in an area rely on system 1 or 2 thinking for true and false statements for the same area respectively, I will look at systematic differences in the time spent to judge the truth status of statements.

When using system 1 thinking, subjects should on average spend less time submitting their answers. System 2 thinking, however, should invite longer time spent on answering. Thus hypotheses 5 and 6 cover the testing of the prior assumptions.

Hypothesis 5: "Subjects with high familiarity in the relevant field will on average spend more time rating the truth status of false statements from their field than subjects with low familiarity."

Hypothesis 6: "Subjects with high familiarity in the relevant field will on average spend a similar time rating the truth status of true statements from their field as subjects with low familiarity."

Finally, due to my interest in the relevance of overconfidence on the size of the illusory truth, the remaining hypotheses cover general as well as field-specific overconfidence. This means, individuals who show signs of overconfidence over all questions of the knowledge test together, as well as individuals who show signs of overconfidence concerning the economics and medicine parts of the quiz separately.

Concretely, my belief is that overconfident individuals are more likely to rely on fluency to judge the truth status of both true and false statements. Overconfidence in this context is

measured by adding together the numerically converted confidence levels for false answers in the knowledge test. High values correspond to comparatively overconfident tendencies. This is done for all questions together and for economically and medically related questions separately.

Additionally, in this context, being overconfident in one's own performance in the knowledge test can be interpreted as a "self-perceived familiarity". Individuals would be expected to be more confident, and thus also more overconfident, in their answers if they perceive the subject of the question to be more familiar to them. It is thus expected that overconfidence increases the size of the IT-Effect, due to a false sense of familiarity. However, since this form of false familiarity is not based on memory or knowledge, contrary to hypothesis 3, false statements should not trigger system 2 thinking and thus the influence on the IT-Effect should be the same for both true and false statements.

Consequently, the following two hypotheses are set.

Hypothesis 7: "For repeated false and true statements, individuals who show comparatively higher signs of being generally overconfident will see a larger IT-Effect than individuals with comparatively lower signs of overconfidence."

Hypothesis 8: "For field-specific repeated true and false statements, individuals who show comparatively higher signs of field-related overconfidence, will on average see a larger IT-Effect than individuals with comparatively lower signs of field-related overconfidence."

Thus, overall, the findings should indicate more precisely whether knowledge and thus the underlying familiarity with a topic, is beneficial or detrimental to the IT-Effect. Hypotheses 1 and 2 cover the standard test for the Illusory Truth Effect. Hypotheses 3-6 examine the effects of familiarity in a field on the IT-Effect and test one possible mechanism, namely differing thought systems, that could be a driver of said effects. Hypotheses 7 and 8 relate to the role of overconfidence, and thus false familiarity, on the IT-Effect overall and for specific fields.

Section III – Methods

III.1 – Aims of the experiment

To test the hypotheses and answer the research question, an online experiment was conducted. In the scope of the study data on subjects' demographics, including their education level, their main field of engagement, and their self-assessed level of knowledge in the said field was gathered. Furthermore, the experiment included a knowledge test, to get an objective measure of individuals' familiarity with the fields of economics and medicine, as well as to get a measure of overconfidence. Finally, the IT-Effect was measured for both general, and field-specific statements of both true and false validity.

The choice of focusing on medicine and economics as the two specified fields was twofold. First, fake news from the medical field has gained importance in recent years, and is thus of particular interest for research. Second, the main share of participants in the study was expected to have a background in economics, thus choosing this field as the second area of specialisation ensured a sufficient number of familiar subjects. Further, looking at two different fields allows for better generalizability of the results, and makes it possible to uncover possible inherent differences between fields. Finally, by including statements related to trivia it was expected to make the study more engaging for subjects and to reduce the probability of participants answering randomly due to being uninterested, thus ensuring both a higher completion rate and more reliable answers.

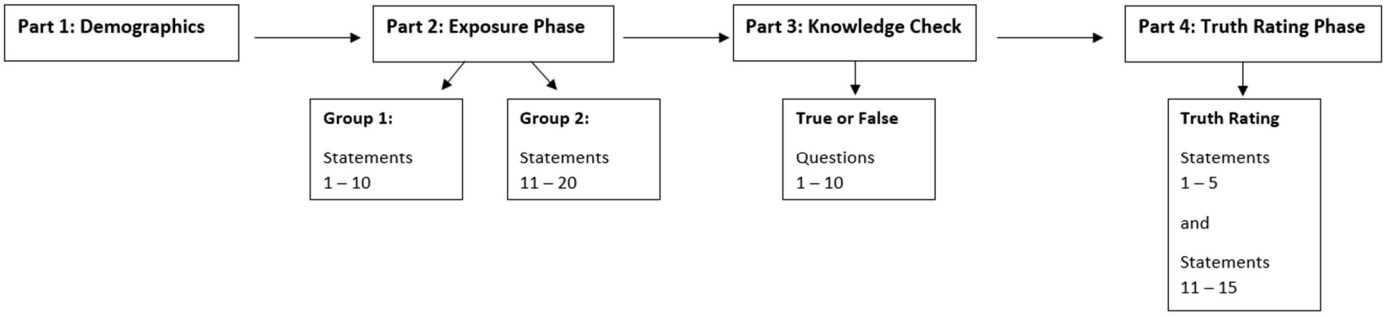
III.2 – Participants

The data collection process started on June the 5th and stopped on August the 3rd 2022. In total, 127 answers were collected, 9 of which were omitted due to incomplete answers or due to showing signs of not having been completed seriously. Thus, the final dataset included responses from 118 subjects. Participants were invited to the study via private connections, text channels from the Erasmus University Rotterdam master's degrees, and survey exchange platforms. Students with an economic or medical background were targeted specifically by inviting them to the study in their respective university text channels. Participants were not provided with any monetary incentives to complete the experiment. Subjects were expected to answer the questions individually, and anonymously without consulting any outside sources.

III.3 – Design and Procedure

The main experiment was structured in four parts (Fig. 3). Due to the similarity of the study and research question, the experiment was closely inspired by previous studies in the field such as Fazio et al. (2015), Boehm (1994), and Srull (1983).

Figure 3: Logical Structure of the Experiment



Note: Subjects in part 2 were randomly assigned to group 1 or 2.

Before starting the experiment, subjects were informed of the guaranteed anonymity of their answers, the general reason for the study, as well as the fact that they would be potentially exposed to fake information during their participation. Individuals had the option to opt out of the study and were provided with the necessary contact information in case of questions or concerns.

In part one of the experiment, subjects first answered questions related to their demographics, such as their age, gender, country of origin, and English proficiency. Subjects also established their level of education, their current status of vocation, the field in which they are either studying or working and their self-assessed level of knowledge in the said field compared to other average individuals active in the same field. Specifically, the question was framed as “How would you rate your level of knowledge in your field compared to the average person active in the same field?”, and subjects had to place themselves on a 7-point scale from “Much below average” to “Much above average” (See Appendix I.1).

In the second part of the experiment, labelled the "Exposure Phase", subjects were randomly assigned to one of two groups, 1 and 2, and according to their group were exposed to a series of ten different statements of varying truthfulness, related to economics, medicine, or trivia. Participants were not forced to respond in a specific time frame, and the order in which the statements appeared was not randomized. Further, while the chosen statements stemmed partially from the same fields, they were chosen so as to not be related to each other or to

questions and statements in the later parts. In total subjects in group 1 read 5 medically related statements of which 4 were false and 1 was true, 3 economically related statements of which 1 was false and 2 were true, and 2 trivia statements of which 1 was false and 1 was true. Subjects in group 2 read 4 medically related statements of which 2 were true and 2 were false, 4 economically related statements of which 3 were false and 1 was true, and two trivia statements of which 1 was false and 1 was true. Both groups thus saw 6 false and 4 true statements in total. Included statements were for example “On average in 2021 women in the EU earned 23% less per hour than men” which is false or “Dementia is currently the seventh leading cause of death among all diseases worldwide” which was true at the time in which the study was conducted (See Appendix I.2). For each statement subjects were asked to rate how interesting the information is on a 6-point scale, ranging from “Very Uninteresting” to “Very Interesting”. This part of the design aimed to ensure that individuals properly read, understood, and somewhat thought about the statements they were presented with. This was further controlled for by measuring the time spent on each question.

In the third part of the experiment, labelled the "Knowledge Test", subjects from both groups were asked to answer 10 true-or-false statements from either economics, medicine, or trivia. Specifically, all subjects answered the same 4 medically related questions, 3 economically related questions, and 3 trivia questions. Additionally, after each answer, individuals were asked to rate how confident they were in their answer on a 6-point scale ranging from “Completely Unconfident” to “Completely Confident”. Included statements were for example “The Solow Growth Model focuses on long-run economic growth” which is true, or “The Kidney produces Insulin” which is false (See Appendix I.3).

The answers from this section of the experiment served as the basis for both the objective level of knowledge in medicine or economics, as well as the comparative level of overconfidence that subjects portrayed either over all statements combined or specifically for those related to medicine or economics. Statements were chosen, as to be relevant enough to allow for a measurement of knowledge, and at the same time to be different enough from the statements used in the second and fourth phases so as to not interfere with, influence, or be influenced by the answers of individuals at any point in time.

In the fourth and last part of the experiment, labelled the "Truth Rating Phase", subjects were again exposed to a series of 10 statements. Individuals from groups 1 and 2 were shown the

same statements. However, half of the statements were drawn from the set of statements group 1 encountered in part two of the study, while the other half were drawn from the set group 2 encountered. Thus, in total, all subjects saw 4 repeated medically related statements of which 3 were false and 1 was true, 4 repeated economically related statements of which 2 were true and 2 were false, and 2 repeated trivia statements of which 1 was true and 1 was false (See Appendix I.4).

The reason for having two groups who encountered different statements in the exposure phase was that this accounts for inherent differences in the truth ratings of statements. By having all statements deliver data on repeated and new truth ratings, the results are comparable and possible differences between statements can be nullified. To establish the truth rating, participants were asked to rate the perceived truthfulness of each statement individually. The exact phrasing was "How truthful do you believe the statement to be?" and the scale used was a 6-point scale, ranging from "Definitely False" to "Definitely True". A neutral option was not included in order to avoid systematic neutral positions whenever the truth status was very difficult to ascertain.

After completing the main part of the experiment, subjects were again informed that some of the statements they encountered were false, and were provided with a link to a document containing clarifications and sources for all potentially encountered pieces of information. This was done in line with the ethical guidelines set up by the Erasmus School of Economics, to avoid any potential lasting negative effects and to guarantee adherence to the scientific ethical standards.

Two distinct choices were made in the design of the experiment. First, since the duration of the delay between the first and second exposure to statements has only been found to have a major effect on the IT-Effect after long periods of time such as days, weeks, or months but not for short periods of time such as a few minutes (Dechêne et al., 2010; Henderson et al., 2021), the knowledge test also acted as a filler task between parts 2 and 4 of the experiment, and was expected to not otherwise influence the results in the truth rating phase. The aim was to make the intentions of the experiment less obvious by not having the truth rating phase directly after the exposure phase and risking subjects realizing a pattern of new versus repeated statements. Second, the low number of statements used for parts 2,3, and 4 compared to previous studies is primarily due to the experimental setting. While lab

experiments are relatively free to take more time without losing subjects along the way, online experiments rely on the willing participation of individuals, and longer studies risk a lower completion rate (Liu and Wronski, 2019). Without any additional incentives for subjects to successfully finish the experiment, the structure of the study had to keep both the attention and interest of participants for as long as possible.

Finally, all statements included in the experiment were chosen after consultation with experts in the respective fields, and taking into account their length, fluency, and difficulty to understand as much as possible. However, no pilot tests were conducted to test the base truth ratings and fluency of all statements over larger samples.

III.4 – Evaluation of data and variables of interest

To be able to properly evaluate the Illusory Truth Effect as well as the influence thereupon of the various mentioned factors, certain variables were converted to a numerical scale. The truth ratings were converted to a numerical scale of 1-6, 1 being "Definitely False" and 6 being "Definitely True", thus a 3.5 would represent uncertainty regarding the truth status of a statement. Further, the confidence scale was converted, 1 being "Completely Unconfident" and 6 being "Completely Confident". This enabled the ability to measure average confidence within and between subjects. Also, answers from the knowledge test were added, with correct answers awarding 1 point and false answers 0. Thus, in total subjects were able to score up to 10 points over all statements, up to 3 points for economic statements, and up to 4 points for medical statements. This delivered the variables for knowledge levels overall, in economics, and in medicine. Consequently, subjects with a score above the median in economics or medicine would be classified as highly familiar with the respective topic, while all others are classified as having low familiarity. A subject could, therefore, be highly familiar in both economics and medicine and vice versa.

To get a measure of relative overconfidence, for each wrong answer in the knowledge test, the corresponding numerical value of confidence was added. Thus, if a subject indicated being “somewhat confident” and gave a wrong answer, his overconfidence score increased by 4. Subjects whose aggregated overconfidence scores were above the median were thus classified as showing comparatively high signs of overconfidence, and all others were classified as showing comparatively low signs of overconfidence.

Since the time spent on each question was measured in the experiment, I was also able to create variables for the average time spent on judging the truth status of false and true, economic and medical statements in part 4 of the experiment.

Finally, I measured the mean truth ratings for all pertinent categories. This included all relevant combinations of false, true, repeated, new, economic, and medical statements. While trivia questions were also a separate category, they were not specifically tested for because they could not be accurately joined into one specific field, as the covered topics are too distant from each other and one could not categorize someone as highly familiar with the subject of trivia given the limited tests in this experiment.

Section IV – Results

All data was analysed and all tests were made using Stata 16.1.

III.4.1 – Descriptive Statistics

Of the 118 participants of the experiment, 76 were Female and 42 were Male. The average age was about 28 years, the youngest being 20. Around 36% of subjects were Dutch and 24% were German, with the remaining subjects stemming from a large variety of countries. All subjects reported at least a decent understanding of English, with 84% reporting a complete understanding or better. 53 subjects indicated to be working or studying in the field of economics and 14 in the field of Medicine or Pharmacy. On the question "How would you rate your level of knowledge in your field compared to the average person active in the same field?", the median answer was "Slightly above average". Thus, suggesting confirmation of previous findings on overconfidence and the so-called better-than-average effect (Alicke and Govorun, 2005). In terms of education, about half the participants claimed to have completed a Bachelor's degree or comparable, a quarter stated to have completed a Master's degree and the remainder was split between having achieved secondary education, a Ph.D., and other unspecified levels of education.

The data from subjects in groups 1 and 2 were pooled together, as no systematic differences between the two were found. This was established by performing t-tests on the mean truth ratings for true and false statements by group, on the mean results in the knowledge test, on the mean age, and on the mean time spent on each statement in part 4 of the experiment. None of the tests showed significant differences between the groups, suggesting that the randomization was successful. Furthermore, when testing for differences between groups for

repeated and new statements, systematic differences were found, suggesting that the statements used for the experiment may have inherent differences in base validity, fluency, or plausibility, emphasizing that the experimental design with two groups was necessary to ensure unbiased results. In other words, instead of comparing a repeated statement with a different new statement, the two-group setup allows a comparison of the same statement being repeated or new.

The following were the mean truth ratings (Ranging from 1-6) of the various defined relevant categories.

Over all categories combined:
Repeated true statements had an average rating of 4.49 (SE [1]= 0.78). New true statements had an average rating of 4.30 (SE = 0.81). Repeated false statements had an average rating of 3.64 (SE = 1.13). New false statements had an average rating of 3.31 (SE = 1.13).

Over economically related statements:
Repeated true statements had an average rating of 5.03 (SE = 1.04). New true statements had an average rating of 5.18 (SE = 0.90). Repeated false statements had an average rating of 4.05 (SE = 1.42). New false statements had an average rating of 3.62 (SE = 1.65).

Over medically related statements:
Repeated true statements had an average rating of 4.16 (SE = 0.97). New true statements had an average rating of 3.75 (SE = 1.17). Repeated false statements had an average rating of 3.48 (SE = 1.10). New false statements had an average rating of 3.07 (SE = 1.21).

Regarding the knowledge test, the median overall score was 6 (Out of 10), the median economics score was 2 (Out of 3), and the median medicine score was 2 (Out of 4). Of the subjects indicating to be studying or working in the economic field, 25% scored less than 2, and only 23% had a full score in economics. Of the subjects indicating to be studying or working in the medical field, all scored at least 3 and 71% had a full score in medicine. Consequently, based purely on the median results from the knowledge test, 53 subjects were classified as having high familiarity with the medical sector, and 65 were classified as being

[1] SE = Standard Error

less familiar with the medical sector. For the economic sector, 24 subjects were placed in the high familiarity category and 94 in the low familiarity category.

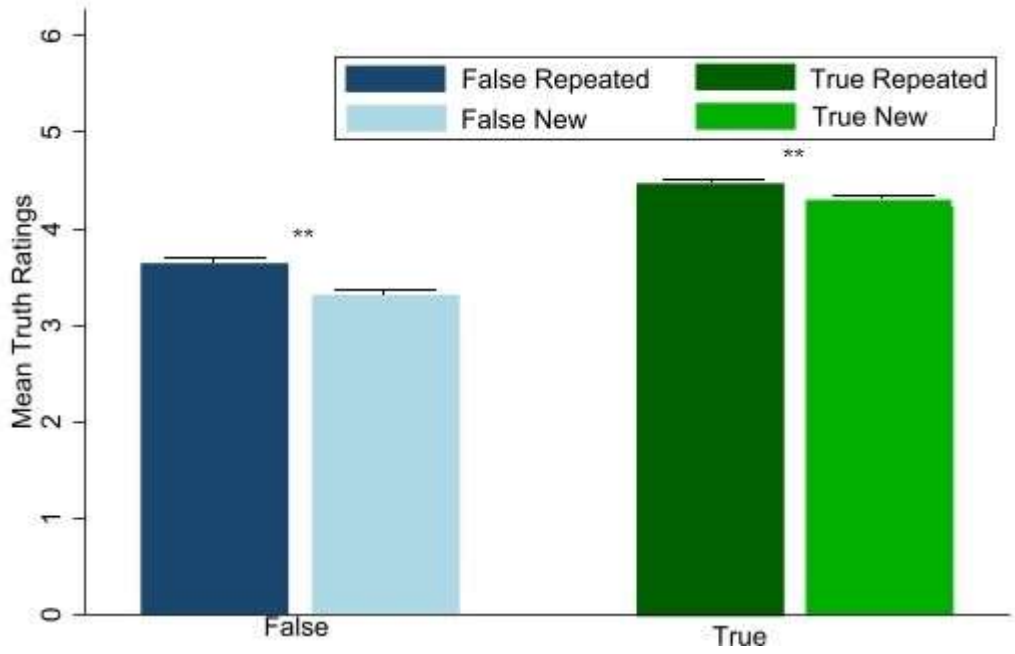
As for the overconfidence scores, the median overconfidence score overall was 12, thus classifying 56 subjects as having comparatively high overconfidence, and 62 as having comparatively low overconfidence. The median overconfidence score for economics was 2, thus classifying 54 subjects as having comparatively high overconfidence and 64 subjects as having comparatively low overconfidence in economics. The median overconfidence score for medicine was 4, thus classifying 52 subjects as having comparatively high overconfidence and 66 as having comparatively low overconfidence in medicine.

Finally, the gathered information on the average time spent on each statement in the truth rating phase showed the following. At the median average 6.66 seconds were spent on each statement overall. For economic statements, the median average time was 6.29 overall, 6.67 for false, and 5.92 for true statements. For medical statements, the median average time was 6.73 overall, 6.83 for false, and 6.12 for true statements.

III.4.2 – General Illusory Truth Effect

To test the IT-Effect I conducted t-tests, comparing the means of the truth ratings between the relevant groups. A t-test allows one to test for statistical differences in group means between two variables. In this case, the relevant opposed variables are, repeated and new statements, repeated and new false statements, and repeated and new true statements. Figure 4 shows a visual representation of the Illusory Truth Effect, clearly showing systematically higher truth ratings for repeated over new statements in general, regardless of them being false or true.

Figure 4: Mean Truth Ratings of Repeated and New False and True Statements



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: False Repeated: 3.64 (SE = 0.10); False New: 3.31 (SE = 0.10); True Repeated: 4.49 (SE = 0.07); True New: 4.30 (SE = 0.07). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

The t-tests were used to evaluate whether the means between the relevant categories were equal. I found that the mean truth rating for repeated false statements was statistically significantly higher than that of new false statements, at a 5% significance level, showing a difference of 0.34 (Repeated: Mean = 3.64; New: Mean = 3.31; $t(117) = 1.87$; $p<0.03$). For repeated and new true statements, the mean for repeated true statements was also statistically significantly higher than that of new true statements, at a 5% significance level, showing a difference of 0.19 (Repeated: Mean = 4.49; New: Mean = 4.30; $t(117) = 1.77$; $p<0.04$).

Therefore, in line with previous research in this area, the collected data supports the idea of a general IT-Effect for both true and false repeated statements. The t-tests, thereby confirm hypotheses 1 and 2. Repetition of statements was found to increase the average truth ratings for both true and false statements. This finding comes as no big surprise, but it helps to confirm previous research and opens up the investigation of further questions.

III.4.3 – Field-specific Illusory Truth Effect

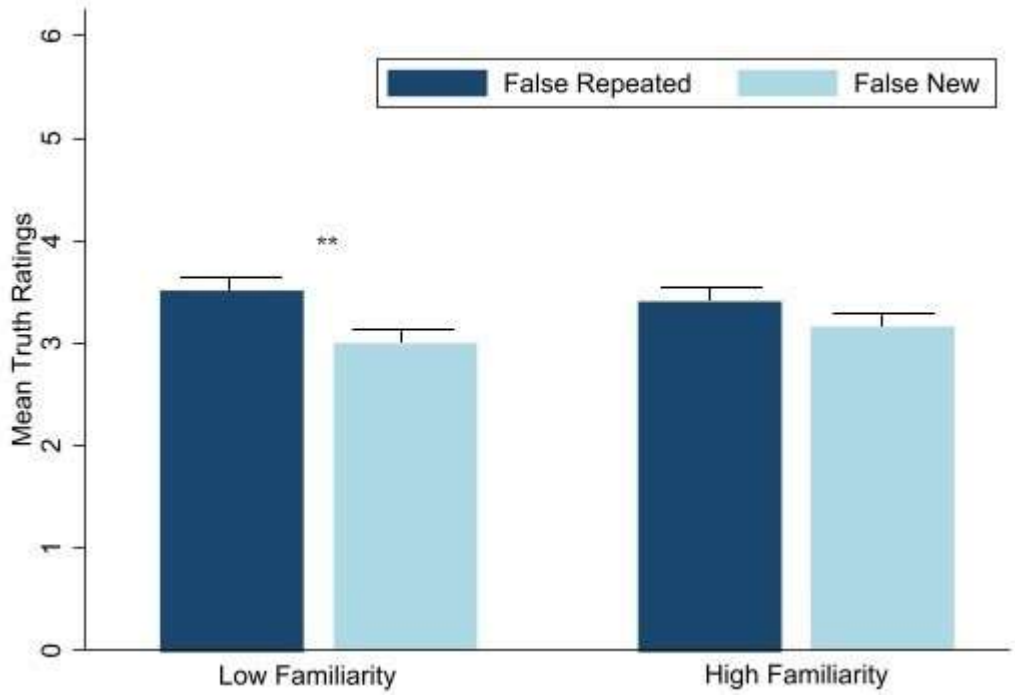
The specific fields covered by my study were economics and medicine. Again t-tests were used to compare the mean truth ratings between groups.

Medicine:

All results presented in this section are in relation to repeated and new medical statements.

The 53 Individuals displaying a high familiarity with the medical sector showed no statistically significant IT-Effect when comparing repeated and new false statements (Repeated: Mean = 3.42; New: Mean = 3.16; $t(52) = 0.98$; $p<0.16$). In comparison, the 65 individuals who were classified as less familiar with the medical sector showed a statistically significantly higher mean truth rating, at the 5% significance level, for repeated false statements than for new false statements, showing a difference of 0.53 (Repeated: Mean = 3.53; New: Mean = 3; $t(64) = 2.35$; $p<0.011$) (See Fig. 5).

Figure 5: Mean Truth Ratings for False Medical Statements by Familiarity

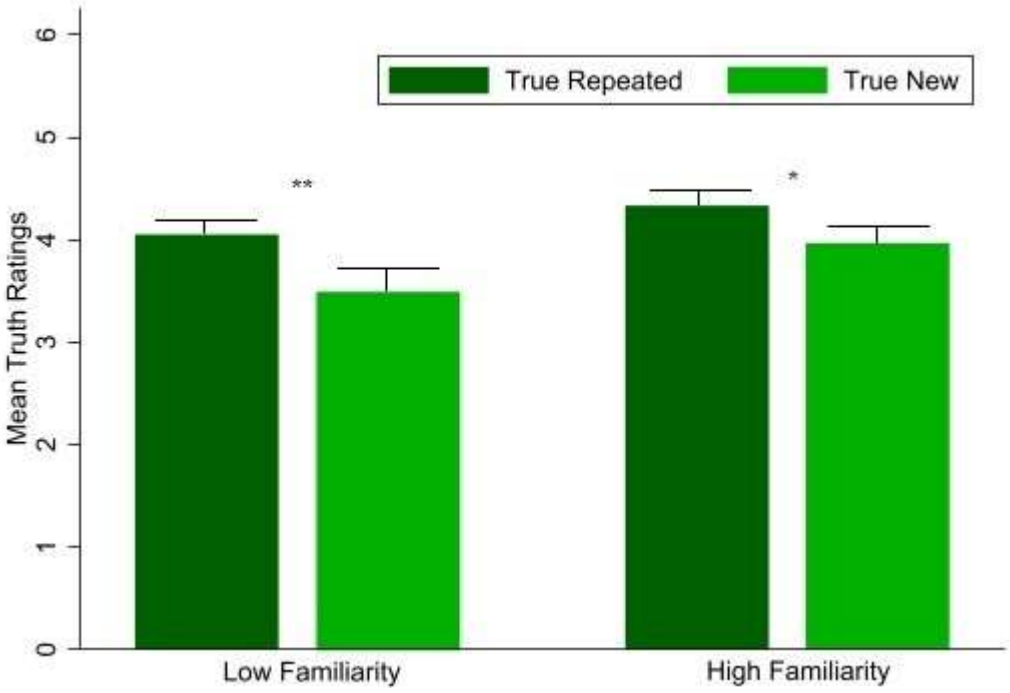


Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low familiarity subjects: False Repeated: 3.53 (SE = 0.14); False New: 3 (SE = 0.14). For high familiarity subjects: False Repeated: 3.42 (SE = 0.15); False New: 3.16 (SE = 0.18). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

For true repeated and new statements, an unpaired t-test had to be performed due to the truth rating phase only including one true medical statement. The test showed that the 53

subjects with high familiarity in the medical sector rated repeated true statements as statistically significantly more truthful than new true statements, at a 10% significance level, resulting in a difference of 0.37 (Repeated: Mean = 4.33; New: Mean = 3.97; t(51) = 1.37; p<0.09). The 65 individuals categorized as less familiar with the medical sector displayed a numerically larger statistically significant IT-Effect for true statements with a difference of 0.55, at a 5% significance level (Repeated: Mean = 4.05; New: Mean = 3.5; t(63) = 1.93; p<0.03)(See Fig. 6).

Figure 6: Mean Truth Ratings for True Medical Statements by Familiarity



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low familiarity subjects: True Repeated: 4.05 (SE = 0.16); True New: 3.5 (SE = 0.25). For high familiarity subjects: True Repeated: 4.33 (SE = 0.18); True New: 3.97 (SE = 0.2). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

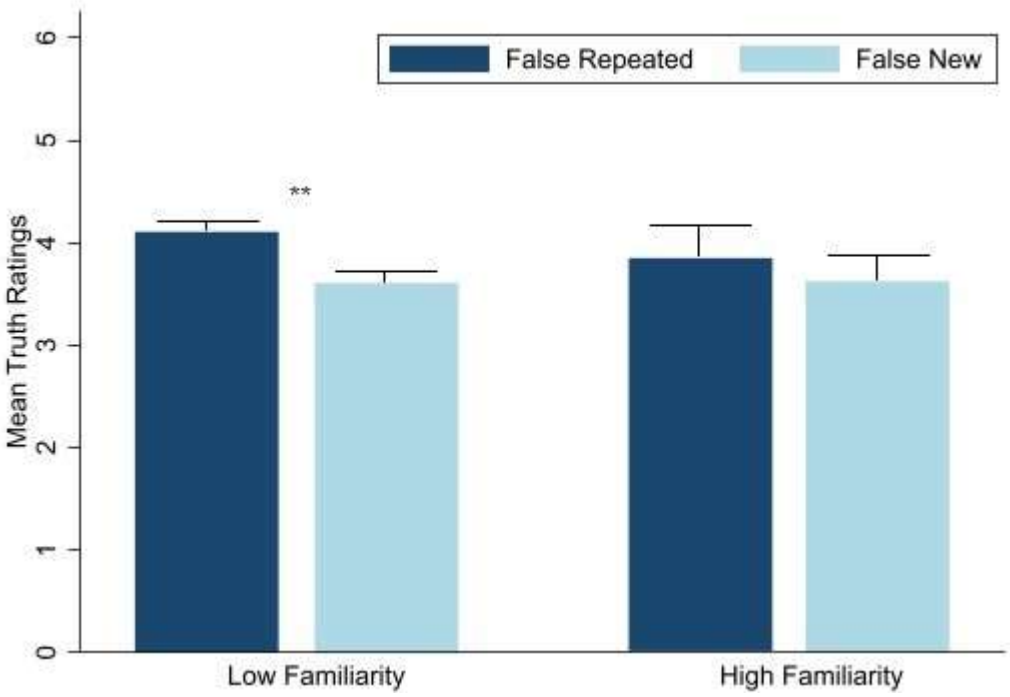
Economics:

All results presented in this section are in relation to repeated and new economic statements.

The 24 individuals classified as displaying a higher familiarity with economics showed no statistically significant IT-Effect for false statements (Repeated: Mean = 3.83; New: Mean = 3.63; t(23) = 0.42; p<0.34). The 94 individuals said to have low familiarity with the economic sector showed a statistically significant IT-Effect at the 5% significance level. The difference in

mean truth rating between repeated and new false statements was 0.49 (Repeated: Mean = 4.11; New: Mean = 3.62; t(93) = 1.86; p<0.03) (See Fig. 7).

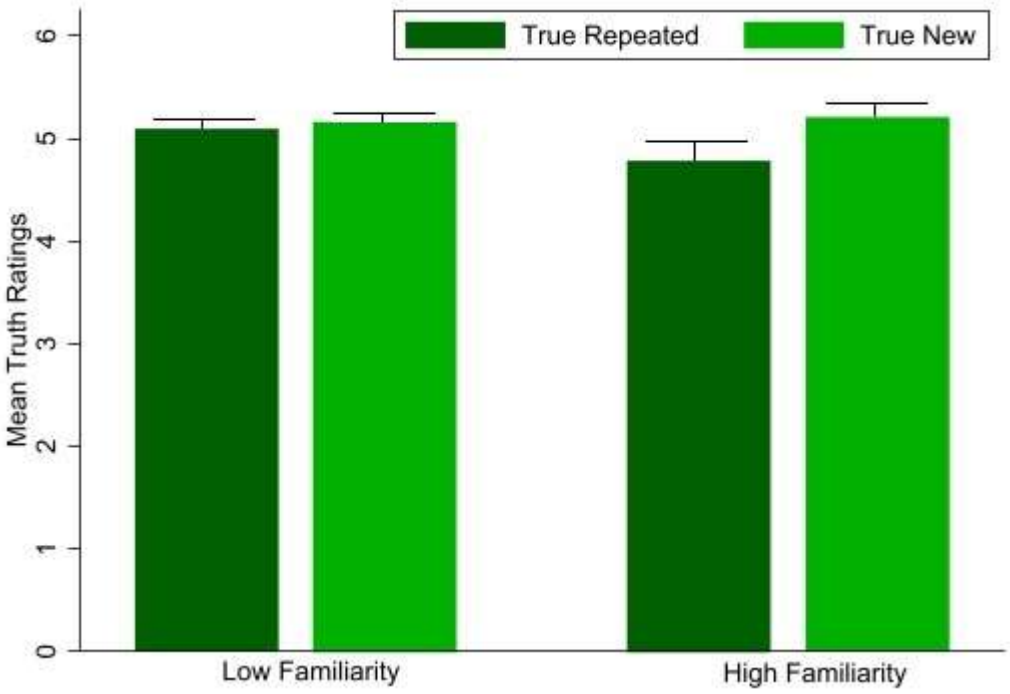
Figure 7: Mean Truth Ratings for False Economic Statements by Familiarity



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low familiarity subjects: False Repeated: 4.11 (SE = 0.14); False New: 3.62 (SE = 0.18). For high familiarity subjects: False Repeated: 3.83 (SE = 0.32); False New: 3.63 (SE = 0.3). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

For true statements, the 24 individuals with high economic familiarity showed no statistically significant IT-Effect. (Repeated: Mean = 4.79; New: Mean = 5.21; t(23) = -1.39; p<0.91). The 94 low familiarity individuals also showed no significant IT-Effect (Repeated: Mean = 5.1; New: Mean = 5.17; t(93) = -0.49; p<0.69) (See Fig. 8). The relatively high overall mean truth ratings for both repeated and new true economic statements may imply that ceiling effects limited the observability of an IT-Effect, as has been the case in previous studies (Fazio et al., 2015). Dechêne et al. (2010, p.254) describe it as "a statement cannot be 'truer' than 'definitely true' by definition", emphasizing the difficulty of observing an IT-Effect when statements' truth status is inherently highly rated.

Figure 8: Mean Truth Ratings for True Economic Statements by Familiarity



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low familiarity subjects: True Repeated: 5.1 (SE = 0.11); True New: 5.17 (SE = 0.1). For high familiarity subjects: True Repeated: 4.79 (SE = 0.21); True New: 5.21 (SE = 0.16). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

Summary:

In summary, regarding the impact of field-related knowledge on the IT-Effect, the results seem to confirm hypothesis 3, namely that a high familiarity with a topic will reduce the IT-Effect for false statements from said topic. Both for economic statements and medical statements evidence of an IT-Effect was only observed in low familiarity individuals of the respective fields. In regards to hypothesis 4, namely that high and low familiarity individuals will see an increased IT-Effect for true statements, the results are not entirely conclusive. However, they suggest that the hypothesis may not hold true. In the case of medical statements, both high familiarity and low familiarity subjects displayed a significant IT-Effect. However, a numerically larger effect size was found for low familiarity subjects than for high familiarity subjects. In the case of economic statements, the IT-Effect was statistically insignificant for both high and low familiarity individuals. One possible explanation that may have limited the observation of an IT-Effect in the case of true economic statements is a ceiling effect. This means that statements were too known, or the truth status too easy to deduct, resulting in overall near

maximum truth ratings and thus leaving little room for observable repetition effects. A conclusive answer to hypothesis 4 may therefore overstate the scope of the results.

III.4.4 – The Relation Between Time and Familiarity in the Truth Rating Phase

To help the argumentation surrounding hypotheses 3 and 4, hypotheses 5 and 6 evaluate the assumption that false statements trigger system 2 thinking in high familiarity subjects and that true statements do not trigger system 2 thinking in high familiarity subjects. As a proxy for system 1 and 2 thinking, the higher the time spent on a statement in the truth rating phase the more conscious the thought process is considered to be. Thus, a high processing time is considered indicative of system 2 thinking, while quick answers are considered indicative of system 1 thinking (Kahneman, 2011).

To test these hypotheses, I first performed t-tests to evaluate whether the average time spent on false or true statements from a certain field was different between high and low familiarity subjects in the respective fields. In a second step, as robustness checks, I performed simple linear regressions to evaluate the relationship between time spent on false statements and the level of familiarity in the relevant area, also taking into account individual characteristics not covered by the simple t-test (See Appendix, II.7).

All 4 t-tests regarding both true and false, medical and economic statements showed no evidence of any statistically significant differences in the time spent on the truth rating between high and low familiarity subjects in the respective fields.

Regarding the linear regressions, the dependent variable was the time spent on the truth rating of the respective statement categories (False and True, Economic and Medical). The explanatory variables were a dummy for having high or low familiarity with the relevant subject, a variable indicating the age of subjects, a variable indicating the English proficiency of subjects, and a dummy for the gender of an individual.

For true and false medical statements, no statistically significant correlation between the time spent on the truth rating and familiarity was found. Instead, most of the differences in time spent on the truth rating seemed to be explained by the age of the subjects. Age showed a statistically significant positive effect on time spent, at a 1% significance level for false medical statements, and a 5% significance level for true medical statements. Thus, indicating that older individuals spent more time on the truth rating.

For false economic statements, a negative correlation between high familiarity and the time spent on the truth rating was found. The effect was significant at the 10% significance level and indicated that subjects with high familiarity in Economics spend on average 2.8 seconds less on the truth rating of false economic statements than subjects with low familiarity, ceteris paribus. For true economic statements, no significant correlation between the time spent on the truth rating and familiarity was found. Again, for both true and false statements, age seemed to be statistically significantly positively correlated with the time spent on the truth rating of statements. In both regressions, the effect of age was significant at a 5% significance level.

These findings suggest a series of possible implications. First, assuming that the assumption still holds that false statements trigger system 2 thinking in individuals with high familiarity in the relevant field, given that there are no apparent systematic differences in the time spent on the truth rating, time may not in fact be a perfect indicator for system 2 or 1 thinking. Second, assuming that the previous assumption does not hold, the findings regarding hypotheses 3 and 4, would need to be explained by a factor of influence other than a difference in thought systems triggered by false statements.

Consequently, the results cannot confirm hypotheses 5 and 6.

III.4.5 – Overconfidence and the Illusory Truth Effect

The last factor of interest that could potentially influence the size of the IT-Effect is false familiarity caused by overconfidence in a matter. As mentioned previously, overconfidence can arise and be especially high when subjects are new to a field (Sanchez and Dunning, 2018). It would therefore be of particular importance to explore whether overconfidence in general or in a specific field does indeed worsen the IT-Effect, and make people more susceptible to it.

General:

All results presented in this section are in relation to the entire set of repeated and new statements.

For false repeated and new statements, the 56 individuals said to be comparatively more overconfident showed an IT-Effect that was statistically significant at the 5% significance level. The difference in mean truth ratings was 0.47 (Repeated: Mean = 3.74; New: Mean = 3.27;

t(55) = 1.78; p<0.04). For the 62 subjects who were deemed comparatively less overconfident, the IT-Effect did not appear significant (Repeated: Mean = 3.55; New: Mean = 3.34; t(61) = 0.87; p<0.19) (See Appendix, II.1).

The 56 Individuals who were classified as comparatively more overconfident were found to have a difference in means between repeated and new true statements of 0.24, which was statistically significant at a 10% significance level (Repeated: Mean = 4.52; New: Mean = 4.28; t(55) = 1.50; p<0.07). The 62 individuals who were classified as comparatively less overconfident displayed no significant IT-Effect for true statements (Repeated: Mean = 4.47; New: Mean = 4.32; t(61) = 0.99; p<0.16) (See Appendix, II.2).

Medicine:

All results presented in this section are in relation to repeated and new medical statements.

Regarding false medical statements, the 52 subjects classed as comparatively more overconfident in medicine, displayed an IT-Effect that was statistically significant at the 5% significance level. The difference between truth ratings of new and repeated statements was 0.413 (Repeated: Mean = 3.49; New: Mean = 3.08; t(51) = 1.69; p<0.049). The 66 subjects classed as being comparatively less overconfident in medicine also displayed an IT-Effect that was statistically significant at the 5% significance level. The difference between truth ratings of new and repeated statements was 0.409 (Repeated: Mean = 3.47; New: Mean = 3.07; t(65) = 1.68; p<0.049)(See Appendix, II.3).

In regards to true medical statements, the 52 subjects classed as comparatively more overconfident in medicine, displayed an IT-Effect that was statistically significant at the 5% significance level. The difference between truth ratings of new and repeated statements was 0.52 (Repeated: Mean = 4.17; New: Mean = 3.65; t(50) = 1.68; p<0.049). The 66 subjects classed as being comparatively less overconfident in medicine did not display a significant IT-Effect (Repeated: Mean = 4.15; New: Mean = 3.81; t(64) = 1.29; p<0.101)(See Appendix, II.4).

Economics:

All results presented in this section are in relation to repeated and new economic statements.

Regarding false statements, the 54 Subjects deemed comparatively more overconfident in economics, indicated a statistically significant IT-Effect, at the 5% significance level. The difference in means between repeated and new statements was 0.81 (Repeated: Mean = 4.17;

New: Mean = 3.35; t(53) = 2.09; p<0.02). The 64 subjects who were classed as comparatively less overconfident in economics, indicated no significant IT-Effect (Repeated: Mean = 3.95; New: Mean = 3.84; t(63) = 0.41; p<0.34)(See Appendix, II.5).

For true statements, neither the 54 subjects classed as comparatively more overconfident in economics (Repeated: Mean = 5.09; New: Mean = 5.2; t(53) = -0.55; p<0.71) nor the 64 subjects classed as comparatively less overconfident displayed a significant IT-Effect (Repeated: Mean = 4.98; New: Mean = 5.16; t(63) = -0.92; p<0.82)(See Appendix, II.6). Again, truth ratings were close to the maximum, indicating possible ceiling effects limiting the effect of repetition.

Summary:

Overall, the findings seem to confirm hypothesis 7, namely that individuals deemed comparatively more overconfident overall will display a larger IT-Effect, regarding both true and false statements overall, than those deemed comparatively less overconfident.

In regards to hypothesis 8, however, the results seem inconclusive. The hypothesis states that for individuals deemed comparatively more overconfident in a field, the IT-Effect for true and false statements should be larger than for those individuals classed as comparatively less overconfident. For true medical statements and false economic statements such an effect was observed, however, for false medical statements the IT-Effect for both groups was of a similar size, and for true economic statements there was no evidence of an IT-Effect for the two groups.

Section V – Analysis

Summarizing all the results found in Section IV, several implications follow.

The findings suggest that a general IT-Effect is present in the analysed dataset, suggesting that the base experimental setup was successful. As for the effects of high familiarity with a subject, determined by a knowledge test, it appears that high familiarity in a field counters the effects of repetition for false but not for true statements from the said field.

In regards to the previous findings by Srull (1983), Arkes et al. (1989), and Boehm (1994) these results are, therefore, novel. Srull (1983) finds that repetition effects appear less for high familiarity than low familiarity subjects, while Arkes et al. (1989) find the IT-Effect to disappear

with low familiarity and increase with high familiarity. However, neither of the two studies differentiated between the IT-Effect for false and true statements. Boehm (1994) found that an increase in familiarity increased the IT-Effect for both true and false statements. Consequently, when looking at the three prior studies and the current findings from this study, the conclusion to take seems uncertain. It is thus important to emphasize one major factor that could influence the results and conclusions of all 4 studies. This factor is the definition and categorization of familiarity in a field.

Srull (1983), similar to my study, relies on an objective test, the results of which categorize subjects into a high or low familiarity category. Arkes et al. (1989) let subjects self-report their level of familiarity with a topic and assign them to high and low familiarity categories based on this more subjective measure. Boehm (1994) argues that the field of study of subjects is an objective indicator of high familiarity and categorizes the participants accordingly. I would argue that of the three mentioned studies only Srull (1983) properly categorizes subjects into high and low familiarity based on an objective measure of knowledge in the respective fields. The categorization by Arkes et al. (1989) is prone to biases, and while it likely captures an effect of perceived familiarity with a topic, it does arguably not capture the relation between actual familiarity in a field and the IT-Effect. In a sense, it is more similar to the relation between overconfidence in a field and the IT-Effect that was also examined in this study. Finally, regarding the categorization by Boehm (1994), in my study working or studying in the field of medicine was an accurate indicator of high familiarity with medical statements, whereas working or studying in the field of economics was not an accurate indicator of high familiarity with economic statements. This means that it is likely that the categorization by Boehm (1994) classifies some individuals as highly familiar with a topic, even though they wouldn't fall under that category based on objective tests. Consequently, this argumentation does not aim to invalidate the previous studies' findings, however, it emphasizes a need for a more specified definition of high familiarity, based on knowledge, in order to replicate and more precisely test the effects on the Illusory Truth Effect.

The results regarding hypotheses 7 and 8 further emphasize this need for standardization. The confirmation of hypothesis 7 shows that in a more general sense, being overconfident may lead to an increased IT-Effect. Further, while my findings are inconclusive regarding the effects of overconfidence in a specific field on true and false statements from the same respective

field, they add to the previous arguments. The IT-Effect found to be driven by overconfidence in a field was seemingly present in some but not all cases, however, more importantly there was no case of overconfidence reducing the IT-Effect. Thus, while inconclusive, the results suggest that the effects of overconfidence found in the general case may potentially translate to the field-specific case. Returning to the study by Arkes et al. (1989), it is thus possible that the findings were not necessarily indicating a relation between actual familiarity with a topic on the IT-Effect, but instead conveyed the influence of false familiarity on the effects of repetition in truth judgements.

Furthermore, regarding the assumption that the effects described in hypotheses 3 and 4 would be driven by a mechanism in which false statements would trigger system 2 thinking in subjects deemed highly familiar with a topic, the findings could not confirm the assumptions. While the findings do not necessarily deny the idea that system 2 rather than system 1 thinking drives the difference in IT-Effects between high and low familiarity subjects, they may suggest that perhaps some other mechanism may be the primary driver. For example, the driving factor could be a difference in neural processes regarding memory that is triggered by encountering false statements in highly familiar topics and that is not actualized via system 2 thinking but rather remains subconscious. The analysis of such a possible relation could thus be the potential topic of future research.

Section VI – Discussion

Summary:
Summarizing all the discussed findings, the IT-Effect seems to be affected to some degree by both actual and false familiarity, as previously defined. The effects of actual familiarity, based on a measure of knowledge in a field, appear rather conclusive, with high familiarity in an area reducing the IT-Effect concerning false statements from said field, but not influencing the IT-Effect concerning true statements from the same field. This finding could not be sufficiently explained by the proposed model of false statements triggering a system 2 thought process, thus potentially suggesting a different mechanism driving these results. The effects of false familiarity, based on a measure of comparative overconfidence in a field, suggest that said false familiarity strengthens the effects of the IT-Effect, or at the very least does not reduce the size of the IT-Effect. However, the conclusions regarding this effect of false familiarity, or

overconfidence, on the Illusory Truth Effect are somewhat inconclusive when examining the case of overconfidence in specific fields.

Thus, In regards to the initial research question, "Does familiarity within a field increase the size of the Illusory Truth Effect for statements from said field, and is overconfidence in the said field a factor that can explain this phenomenon?" the findings suggest that familiarity in a field does not increase the IT-Effect for any statements from said field, and in fact, in the case of false statements from a field it actually seems to decrease the size of the truth effect. As for the role of overconfidence, a comparatively high level of overconfidence seems to invite a stronger IT-Effect than a low level of overconfidence. However, the results were somewhat inconclusive for overconfidence in a specific field. This may be due to the combined effects of true familiarity, based on knowledge, and false familiarity, based on overconfidence. Seeing as some individuals that are perceived to be overconfident may actually have a decent level of knowledge, and thus also a higher true familiarity, the effects of knowledge may somewhat counter the effects of overconfidence when determining the size of the IT-Effect.

Limitations:

Before going into the implications of the findings, and making suggestions for policymakers and future research, it is necessary to address the shortcomings of both the analysis of the collected data as well as the experiment itself.

First, concerning the structure of the experiment itself, some issues arose that were not clear to become a problem before the analysis of the data. In hindsight, focusing on a single specified field of interest such as medicine might have led to an improvement of the analysis. This is due to the fact that by limiting the number of statements to 10 in parts 2, 3, and 4, a smaller sample for the IT-Effect and knowledge levels of subjects in each field of interest were collected. Having 10 statements in each part related to a single field could potentially have delivered more valid results due to larger sample size. Thus, by trying to increase the completion rate of the experiment, the collected data might have gone down in quality.

Additionally, more attention should have been paid to the balancing of statements included in parts 2, 3, and 4. Ideally, in the exposure phase (part 2) subjects in groups 1 and 2 should have been exposed to the same number of economic, medical, and trivia statements. Instead, group 1 saw 5 medical statements and group 2 saw 4, and group 1 saw 3 economic statements

while group 2 saw 4. While this may not have necessarily influenced the results, it should have been balanced to avoid systematic differences between the two groups. Further, regarding the knowledge test (part 3), the same balancing issue arose. While participants answered 4 medical questions, they were only asked to answer 3 economic questions. This may have led to a less accurate categorization of subjects into the groups of high and low familiarity and overconfidence, reducing the comparability of results between medicine and economics. Finally, regarding the truth rating phase (part 4), again balancing was an issue. While subjects rated the truth status of an equal number of economic and medical statements, only the economic statements were balanced between true and false. The medical statements included 3 falsities and 1 truth, while the economic statements included 2 falsities and 2 truths. This again potentially reduces the comparability between findings regarding medicine and economics, and the fact of only having sampled 1 true medical statement may have reduced the validity of the analysis regarding the said statement, due to a low sample size for the IT-Effect.

Furthermore, the statements' base perceived validities were not tested through a pilot study beforehand. This means that some statements' truth status could have been too obvious regardless of familiarity with the field the statements belong to. The existence of this issue is suggested by the high overall truth ratings received by true economic statements, which possibly invited ceiling effects that reduced the ability to test for an IT-Effect.

Fourth, due to the ethical constraint of having to inform subjects of potentially false information, the overall size of the IT-Effect across all mentioned categories may be smaller than in a setting without such warnings (Jalbert et al., 2020; Calio et al., 2020; Nadarevic and Aßfald, 2017). In the study by Jalbert et al. (2020), experiments researching the IT-Effect were performed with and without prior warnings to the subjects that false information may be encountered. The results of these experiments showed that the IT-Effect was much larger when subjects were not warned prior to the truth ratings. In fact, Jalbert et al. (2020) found the IT-Effect to be twice as large for the experiment omitting the warnings regarding false information. Thus, based on these results, the size of the IT-Effects found in my study is likely to be underestimated, compared to a natural setting environment in which no warnings are encountered. Further, this also means that the IT-Effect found in this study may be inherently smaller in size than in studies in which subjects are not warned beforehand.

Furthermore, the mere fact of using a non-field experiment may influence the size of the IT-Effect. As was established previously, the IT-Effect is largely due to differences in fluency and due to the fact that people tend to evaluate claims intuitively. However, one can argue that an experimental setting may invite behaviour that is different from that in the field. In their paper, Falk and Heckman (2009) discuss various criticisms concerning laboratory experiments and argue for the validity and utility of all types of experiments. One important aspect that is discussed is the so-called “Hawthorne effect”. Referring to one of many widely used definitions for this effect (Chiesa and Hobbs, 2008), the theory states that the mere fact of subjects knowingly participating in an experiment, increases their productivity or changes some other tested but technically independent outcome.

While this effect has been widely criticized (Wickström and Bendix, 2000; Kompier, 2006), the inherent argument that subjects behave differently in an experiment than they would otherwise still stands. Individuals may, for example, act more altruistically due to social norms they feel like they should adhere to, due to wanting to please the experimenter, or due to lower stakes than in "real life" (Benz and Meier, 2008). It is, therefore, reasonable to assume that subjects are more thoughtful in an experiment in which they are asked to rate the truthfulness of statements than they would be when doing the same on social media. Participants may feel like they should answer as accurately as possible out of fear of “performing worse” than other participants, or giving “wrong” answers.

In fact, during the duration of the experiment I was approached by several subjects whom I had personally invited to participate in the study, that commented that they felt like they should have known the answers to all questions or that the questions were very difficult. This suggests to me, that at least some portion of the subjects who participated in my study tried especially hard to get accurate results and saw the experiment as a challenge, when in reality, they would not have put as much thought and effort into similar statements encountered on social media. Thus, due to the inherent importance of fluency and the likely differences in the IT-Effect when individuals rely on system 1 or system 2 thinking, the size of the IT-Effect, as well as some of the discussed effects on said size, may have been further underestimated.

Therefore, overall, results should be interpreted with care, and drawn inferences should be made while keeping the potential limitations of the study in mind. However, for further

discussion of the findings, all results will be treated according to their established significance, and predictions and suggestions will be made accordingly.

Implications:

Concluding, the results from the conducted experiment suggest a range of implications. First, seeing as high familiarity with a topic seems to be able to reduce the Illusory Truth Effect for false statements in the relevant field, educative measures that clear up common misconceptions in certain areas and teach the general population on specific subjects may be beneficial to deter the spread of fake news. However, for such measures to be effective the familiarization process needs to be sufficient as leaving individuals with only a little knowledge, and perhaps a false sense of familiarity with a topic, may instead make them more prone to fall prey to the effects of repetition on the credibility of fake news. Second, this thought is further emphasized by the apparent issues of overconfidence present when individuals overestimate their own level of knowledge in a field. Such overconfidence, shown by previous studies to be mostly present in beginners, seems to invite a false sense of familiarity that appears to strengthen the detrimental IT-Effect. Third, this study confirms the findings of one previous study (Srull, 1983), but comes to a somewhat different conclusion than two other studies looking at the effects of knowledge-based familiarity in a field (Arkes et al., 1989; Boehm, 1994). This invites the question of how these differences come to be. As previously suggested the source of this variation may lie in the method of determining high familiarity with a topic, which seems inconsistent across studies. While the experiment conducted in the context of this thesis and the study by Srull (1983) rely on a categorization of familiarity based primarily on knowledge, the studies by Arkes et al. (1989) and Boehm (1994) categorized subjects into high or low familiarity based on self-assessments and using a field of study as a proxy. The two latter methods, however, invite the possibility of including subjects that display a false familiarity, based on overconfidence, rather than a true familiarity, based on knowledge. This means that the results of the two latter studies may be somewhat distorted by the influencing effects of overconfidence when assessing the relation between knowledge and the size of the Illusory Truth Effect.

Furthermore, this study differentiates between high and low levels of familiarity based on knowledge. However, it is likely that the results only depict the aggregated effects, while in reality, the relationship between familiarity and the IT-Effect may be more nuanced. While

experts may be better at not being affected by the repetition of false statements, less familiar individuals, that would still fall under the category of high familiarity in this experiment, may remain affected by an IT-Effect. Consequently, using the fluency-conditional model by Fazio et al. (2015 p.997; Fig.2), one can apply this idea to a theoretical foundation. If one assumes that being active in a field improves the fluency of statements belonging to the said field due to the increased familiarity, subjects who are somewhat knowledgeable may have a stronger tendency to rely on their intuition to judge statements than subjects with no prior knowledge. Up until a person has some related knowledge that strictly and sufficiently contradicts the information contained in the statement at hand, reliance on fluency is thereby likely to increase with knowledge. In essence, it is therefore a difficult task to know just how much knowledge is beneficial and how much is detrimental, both in regards to the IT-Effect and also to fake news in general. The size of the IT-Effect only concerns the effects of repetition on the perception of fake news. However, it disregards the base perception of misinformation. If little knowledge decreases the credibility of false information on the base level and only increases the effects of repetition, then the resulting perception may still be better than without knowledge.

Consequently, in regards to government's approach to fighting fake news, rather than relying fully on uninformative warning labels on items of fake news, informative measures such as education in schools, workplaces, on official social media accounts, and via government-funded unbiased mediums may prove most efficient. However, due to the dangers of increasing people's overconfidence in certain fields, the retention of knowledge should be assured, and thus the information should be provided as clearly as possible. Measures that intend to educate the public on general topics may actually be counter-effective if the transmitted knowledge is either not directly relevant to the fake news, or if it is too broad and individuals end up with little actual knowledge about the subject. Such initiatives should therefore be performed diligently and taking into account the potential drawbacks. I believe that for example during the Covid-19 pandemic, the spread of fake news and the perceived validity of alternative facts were partially aided by governments' approaches to providing information to the public. With governments updating their own beliefs at regular intervals and being too quick to give relatively unfounded health advice, the fluency and retention of truthful information by the public were likely reduced.

Therefore, when confronted with relatively new issues and a rise of misinformation, governments should remain calm and assess the true facts, before providing the public with hasty statements that may need later updating or correcting. While time may be of importance when correcting false beliefs, corrections and the propagation of true and important information should be done in a concise and easy-to-understand fashion, leaving little room for doubt.

Finally, regarding future studies, researchers could concentrate on three potential questions. First, "How much knowledge is enough knowledge when fighting fake news?". This question relates to the variance in the size of the IT-Effect for different levels of knowledge. Researchers should study this variance more, and attempt to establish the different knowledge thresholds that determine the size of the truth effect. Second, “How does the Illusory Truth Effect vary between sectors?”, or “Are Fake News equally efficient across all fields?”. These questions are interesting because it is likely that there are inherent differences in the effectiveness of fake news in general and the Truth Effect specifically when looking at different fields. For example, for the fields used in this study, economically related statements, compared to medically related ones, may have an inherently higher base fluency due to both easier-to-understand language and a more general familiarity with the subject across the entire population. Everyone should understand the concept of wage differences but some may struggle to understand the relationship between dementia and death rates. It may therefore be helpful for future policy-makers to understand the differences between fake news in differing sectors, to be able to adaptively employ the most effective measures. Third, “Are measures against fake news one-size-fits-all?”. Since fake news perception seems to differ between persons of different backgrounds, it would be logical to assume that measures that combat misinformation also vary in effectiveness depending on the individual's characteristics. Studies should therefore be conducted to test the success of different systems against the spread of fake news, in regards to personal characteristics such as familiarity with a field, overconfidence, or even standard demographics such as age, education, and gender.

In conclusion, future policy-makers that aim to combat fake news should take into account these differences in relationships and further explore more specifically employable interventions. Broad all-encompassing measures that “fit all sizes” might not be the right approach to deterring the spread of misinformation. Additionally, future research should

further explore the behavioural effects of knowledge and familiarity, study differences in the IT-Effect between fields, and test the effectiveness of anti-fake-news methods related to different levels of knowledge.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-36.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review*, *13*(3), 219-235.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, *2*(2), 81-94.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human learning and memory*, *5*(3), 241.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*(4), 446.

Begg, I., Armour, V., & Kerr, T. (1985). On believing what we remember. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *17*(3), 199.

Benz, M., & Meier, S. (2008). Do people behave in experiments as in the field?—evidence from donations. *Experimental economics*, *11*(3), 268-281.

Boehm, L. E. (1994). The validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, *20*(3), 285-293.

Brandenberger, D. (2012). Propaganda state in crisis: Soviet ideology, indoctrination, and terror under Stalin, 1927-1941. *Yale University Press*.

Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences*, *118*(5), e2020043118.

Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current directions in psychological science*, *29*(3), 316-323.

Calio, F., Nadarevic, L., & Musch, J. (2020). How explicit warnings reduce the truth effect: A multinomial modeling approach. *Acta Psychologica*, *211*, 103185.

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, *28*(11), 1531-1546.

Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne Effect?. *European Journal of Social Psychology*, *38*(1), 67-74.

De Keersmaecker, J., Dunning, D., Pennycook, G., Rand, D. G., Sanchez, C., Unkelbach, C., & Roets, A. (2020). Investigating the robustness of the illusory truth effect across individual differences in cognitive ability, need for cognitive closure, and cognitive style. *Personality and Social Psychology Bulletin*, *46*(2), 204-215.

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, *14*(2), 238-257.

Doob, L. W. (1950). Goebbels' principles of propaganda. *Public Opinion Quarterly*, *14*(3), 419-442.

Ecker, U. K., Hogan, J. L., & Lewandowsky, S. (2017). Reminders and repetition of misinformation: Helping or hindering its retraction? *Journal of Applied Research in Memory and Cognition*, *6*(2), 185-192.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *science*, *326*(5952), 535-538.

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology*: General, *144*(5), 993.

Fernández-Torres, M. J., Almansa-Martínez, A., & Chamizo-Sánchez, R. (2021). Infodemic and fake news in Spain during the COVID-19 pandemic*. International journal of environmental research and public health*, *18*(4), 1781.

Furini, M., Mirri, S., Montangero, M., & Prandi, C. (2020, July). Untangling between fake-news and truth in social media to understand the Covid-19 Coronavirus. *In 2020 IEEE Symposium on Computers and Communications (ISCC)* (pp. 1-6). IEEE.

Gottfried, J., & Shearer, E. (2016). News use across social media platforms 2016. Pew Research Center (2016).

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of verbal learning and verbal behavior*, *16*(1), 107-112.

Henderson, E. L., Simons, D. J., & Barr, D. J. (2021). The trajectory of truth: A longitudinal study of the illusory truth effect. *Journal of cognition*, *4*(1).

Jalbert, M., Newman, E., & Schwarz, N. (2020). Only half of what I'll tell you is true: Expecting to encounter falsehoods reduces illusory truth. *Journal of Applied Research in Memory and Cognition*, *9*(4), 602-613.

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kompier, M. A. (2006). The" Hawthorne effect" is a myth, but what keeps the story going?. *Scandinavian journal of work, environment & health*, 402-412.

Krieghofer, G. (2017, December 6). "*Eine Lüge muss nur oft genug wiederholt werden. Dann wird sie geglaubt*." falschzitate.blogspot.com. Retrieved August 26, 2022, from https://falschzitate.blogspot.com/2017/12/eine-luge-muss-nur-oft-genug-wiederholt.html

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological science in the public interest*, *13*(3), 106-131.

Liu, M., & Wronski, L. (2018). Examining completion rates in web surveys via over 25,000 real-world surveys. *Social Science Computer Review*, *36*(1), 116-124.

Lyons, B. A., Montgomery, J. M., Guess, A. M., Nyhan, B., & Reifler, J. (2021). Overconfidence in news judgments is associated with false news susceptibility. *Proceedings of the National Academy of Sciences*, *118*(23), e2019527118.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological review*, *115*(2), 502.

Moravec, P., Minas, R., & Dennis, A. R. (2018). Fake news on social media: People believe what they want to believe when it makes no sense at all. *Kelley School of Business research paper*, (18-87).

Moscadelli, A., Albora, G., Biamonte, M. A., Giorgetti, D., Innocenzio, M., Paoli, S., ... & Bonaccorsi, G. (2020). Fake news and Covid-19 in Italy: results of a quantitative observational study. *International journal of environmental research and public health*, *17*(16), 5850.

Mosseri, A. (2016). "*News Feed FYI: Addressing Hoaxes and Fake News*." Newsroom, Facebook, December 15. Retrieved August 28, 2022 from http://newsroom.fb. com/news/2016/12/news-feed-fyi-addressinghoaxes-and-fake-news/.

Nadarevic, L., & Aßfalg, A. (2017). Unveiling the truth: Warnings reduce the repetition-based truth effect. *Psychological Research*, *81*(4), 814-826.

Pan, W., & Hu, T. Y. (2022). More familiar, more credible? Distinguishing two types of familiarity on the truth effect using the drift-diffusion model.

Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., Akhtar, M. S., ... & Chakraborty, T. (2021, February). Fighting an infodemic: Covid-19 fake news dataset. In *International Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation* (pp. 21-29). Springer, Cham.

Pennycook, G., Bear, A., Collins, E. T., & Rand, D. G. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, *66*(11), 4944-4957.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general*, *147*(12), 1865.

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*, *88*(2), 185-200.

Randolph, H. E., & Barreiro, L. B. (2020). Herd immunity: understanding COVID-19. *Immunity*, *52*(5), 737-741.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and cognition*, *8*(3), 338-342.

Sanchez, C., & Dunning, D. (2018). Overconfidence among beginners: Is a little learning a dangerous thing?. *Journal of personality and Social Psychology*, *114*(1), 10.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*(4), 332-348.

Shearer, E., & Mitchell, A. (2021). News use across social media platforms 2020. Pew Research Center (2020).

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*(4), 713-724.

Srull, T. K. (1983). The role of prior knowledge in the acquisition, retention, and use of new information. *ACR North American Advances*.

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of experimental psychology: learning, memory, and cognition*, *43*(12), 1948.

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature* (March 19, 2018).

Unkelbach, C. (2007). Reversing the truth effect: learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 219.

Voigtländer, N., & Voth, H. J. (2015). Nazi indoctrination and anti-Semitic beliefs in Germany. *Proceedings of the National Academy of Sciences*, *112*(26), 7931-7936.

Wahlheim, C. N., Alexander, T. R., & Peske, C. D. (2020). Reminders of everyday misinformation statements can enhance memory for and beliefs in corrections of those statements in the short term. *Psychological Science*, *31*(10), 1325-1339.

Wang, W. C., Brashier, N. M., Wing, E. A., Marsh, E. J., & Cabeza, R. (2016). On known unknowns: Fluency and the neural mechanisms of illusory truth. *Journal of Cognitive Neuroscience*, *28*, 739-746.

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social science & medicine*, *240*, 112552.

Wickström, G., & Bendix, T. (2000). The" Hawthorne effect"—what did the original Hawthorne studies actually show?. *Scandinavian journal of work, environment & health*, 363-367.

Appendix

I – Survey:

I.1 – Part 1

Information

Thank you for participating in this study. My name is Bruno Sonnenschein and this study is part of my Master Thesis in Behavioral Economics.
For any comments, questions or concerns feel free to contact me via e-mail under the following address:

597126bs@eur.nl

In this study you will be exposed to a variety of statements, some of which will include false information. At the end of the study you will be provided with clarifications for all statements you potentially encountered.

All your answers will be recorded anonymously and cannot be traced back to your person.

- [ ] I understand that I may be exposed to false information
- [ ] I willingly participate in this study

How do you identify?

- ○ Male
- ○ Female
- ○ Non-binary / third gender
- ○ Other
- ○ Prefer not to say

What is your age?

What is your country of origin?

What is your highest achieved level of education?

- ○ Primary Education
- ○ Secondary Education
- ○ Bachelors Degree or comparable
- ○ Masters Degree or comparable
- ○ PhD
- ○ Other

Whats your current vocation?

- ○ Student
- ○ Part-Time Employed
- ○ Full-Time Employed
- ○ Retired
- ○ Other

In what field are you mainly active ? (Academic or Professional)

- ○ Economics or Business Administration
- ○ Statistics
- ○ Education
- ○ Law
- ○ Medicine or Pharmacy
- ○ Psychology
- ○ History
- ○ Politics
- ○ Other

How would you rate your level of knowledge in your field compared to the average person active in the same field?

	Much Below Average	Strictly Below Average	Slightly Below Average	Average	Slightly Above Average	Strictly Above Average	Much Above Average
Knowledge-Level	○	○	○	○	○	○	○

How would you rate your level of English Reading proficiency?

	No Understanding	Some Understanding	Decent Understanding	Complete Understanding	Mother tongue
English Reading Proficiency	○	○	○	○	○

Info and Request

For the remainder of the survey take your time to read and answer each question. Please do not at any point in the survey look up information that could influence your answers.

I.2 – Part 2: Exposure Phase

Exposure Phase Intro

In the following section you will be exposed to a series of statements. You will be asked to carefully read each statement and indicate how interesting it is to you.

Exposure Phase Statement 1

Vaccines generally cause the disease

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

About 5.2% of the global population live in extreme poverty

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 2

In 2021 Norway ranked first on the Human Development Index

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

In 2021 Elon Musk was the richest person in the world

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 3

The Kiwifruit originates from China

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

5.3% of all yearly deaths globally result from harmful use of alcohol

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 4

Simultaneous vaccination with multiple vaccines can have adverse effects on the normal childhood immune system

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

On average one in 1000 children worldwide has autism

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 5

1% of adults worldwide suffer from depression

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Obesity does not increase the risk of diabetes

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 6

The Atlantic is the largest ocean on earth

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Pythagoras is known as "The Founder of Geometry"

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 7

In 2021 one US-Dollar was more than one Euro

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Currently the USA embargo 12 different countries

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 8

Restricting access to abortions does not reduce the number of abortions

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Dementia is currently the seventh leading cause of death among all diseases worldwide

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 9

The European Central Bank sets the interest rates for EU member states

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Not all current EU-member-states use the Euro as their currency

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

Exposure Phase Statement 10

Diabetes is the most common chronic disease among children worldwide

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

On average in 2021 women in the EU earned 23% less per hour than men

	Very Uninteresting	Uninteresting	Slightly Uninteresting	Slightly Interesting	Interesting	Very Interesting
How interesting is the statement?	○	○	○	○	○	○

I.3 – Part 3: Knowledge Test

Knowledge Test Intro

In the following section you will be presented with statements. You will be asked to say whether the statements are true or false and to state how confident you are in your answers.

Again, please do not look up any answers.

Knowledge Test Question 1

The Solow Growth Model focuses on long-run economic growth

- ○ TRUE
- ○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 2

The Wealth of Nations was written by Adam Smith

- ○ TRUE
- ○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 3

The enzyme family of cytochromes P450 is of central importance for drug metabolism

- ○ TRUE
- ○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 4

Julius Caesar was the first Roman emperor

- ○ TRUE

○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 5

The Kidney produces Insulin

○ TRUE
○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 6

In 2017 Richard Thaler received the Nobel prize in Economics for his contributions to Game Theory

- ○ TRUE
- ○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 7

Grapefruit does not have a dangerous influence on medicine intake

- ○ TRUE
- ○ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	○	○	○	○	○	○

Knowledge Test Question 8

The tallest mountain in Europe ranks outside the top 25 worldwide

- ◯ TRUE
- ◯ FALSE

How confident are you that you answered correctly

	Completely Unconfident	Very Unconfident	Somewhat Unconfident	Somewhat Confident	Very Confident	Completely Confident
Level of Confidence	◯	◯	◯	◯	◯	◯

Knowledge Test Question 9

Moses took 2 animals of each kind on the ark

- ◯ TRUE
- ◯ FALSE

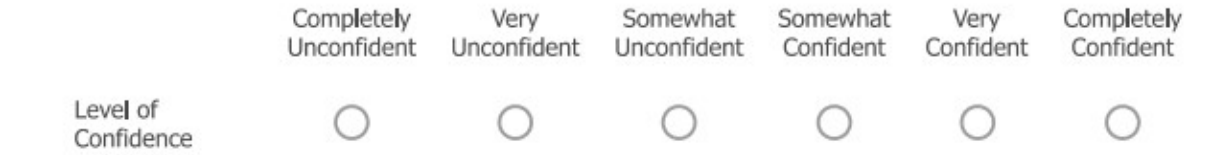
How confident are you that you answered correctly

Knowledge Test Question 10

Suppositories are given to bypass metabolization in the liver

- TRUE
- FALSE

How confident are you that you answered correctly



I.4 – Part 4: Truth Rating Phase

Truth rating phase

In the following section you will be exposed to statements. You will be asked to rate their truthfulness based on your belief.

Truth Rating Phase Statement 1

Vaccines generally cause the disease

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 2

The Kiwifruit originates from China

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 3

1% of adults worldwide suffer from depression

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 4

In 2021 Norway ranked first on the Human Development Index

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 5

In 2021 one US-Dollar was more than one Euro

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 6

On average one in 1000 children worldwide has autism

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 7

Dementia is currently the seventh leading cause of death among all diseases worldwide

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 8

Not all current EU-member-states use the Euro as their currency

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 9

Pythagoras is known as "The Founder of Geometry"

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

Truth Rating Phase Statement 10

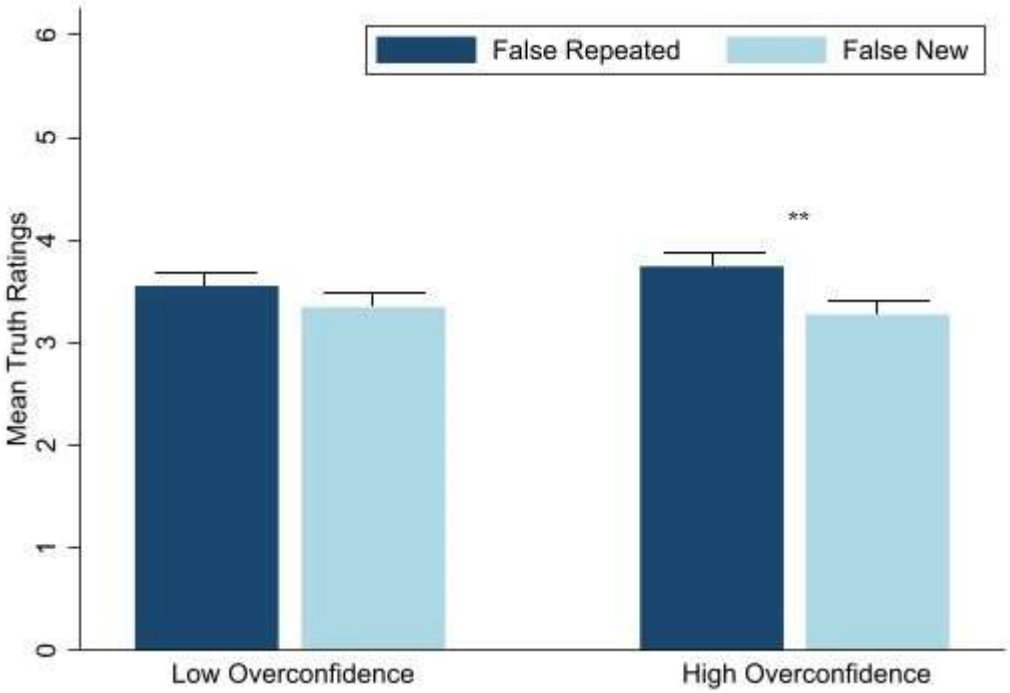
On average in 2021 women in the EU earned 23% less per hour than men

How truthful do you believe the statement to be?

	Definetely False	Probably False	Possibly False	Possibly True	Probably True	Definetely True
Truthfulness	○	○	○	○	○	○

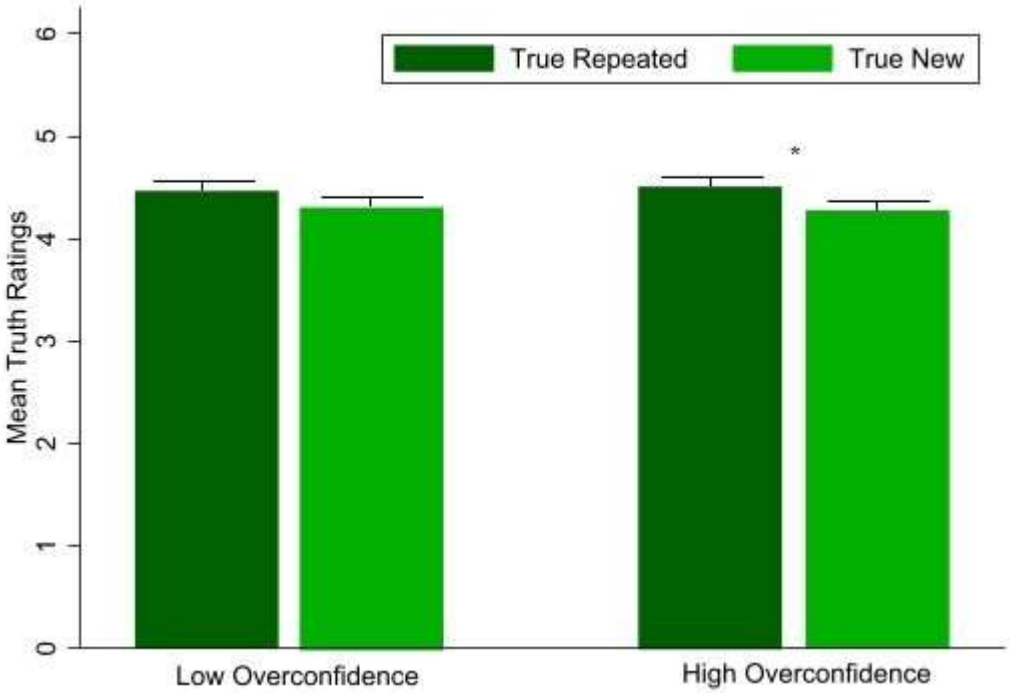
II – Additional Figures:

II.1 - Mean Truth Ratings for False Statements by Overconfidence



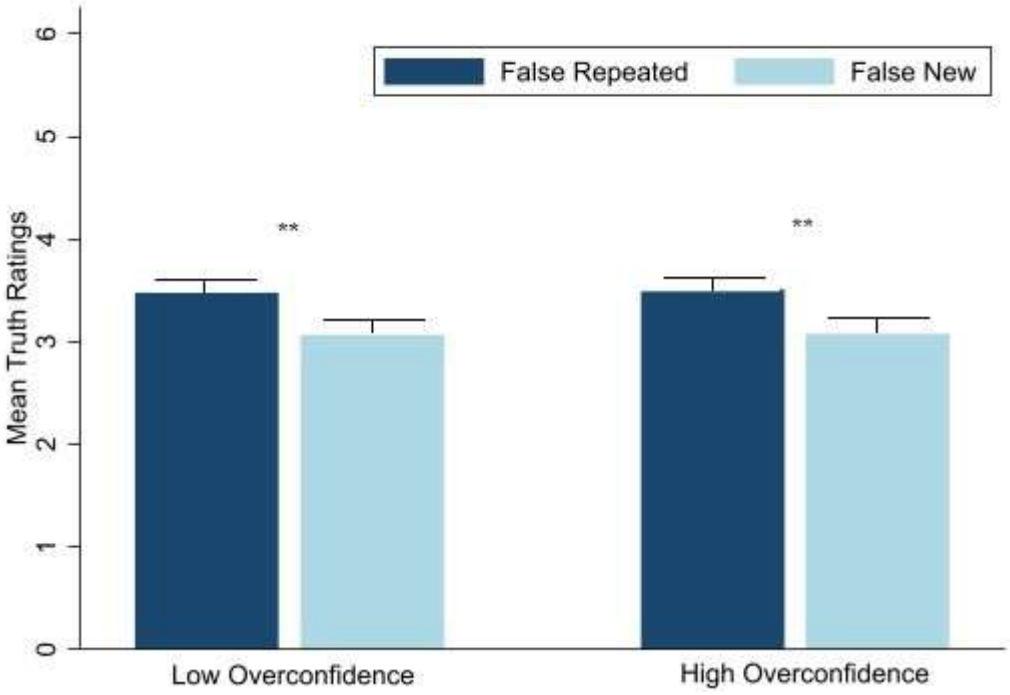
Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: False Repeated = 3.55 (SE = 0.14); False New = 3.34 (0.15). For high overconfidence subjects: False Repeated = 3.74 (SE = 0.16); False New = 3.27 (SE = 0.15). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.2 - Mean Truth Ratings for True Statements by Overconfidence



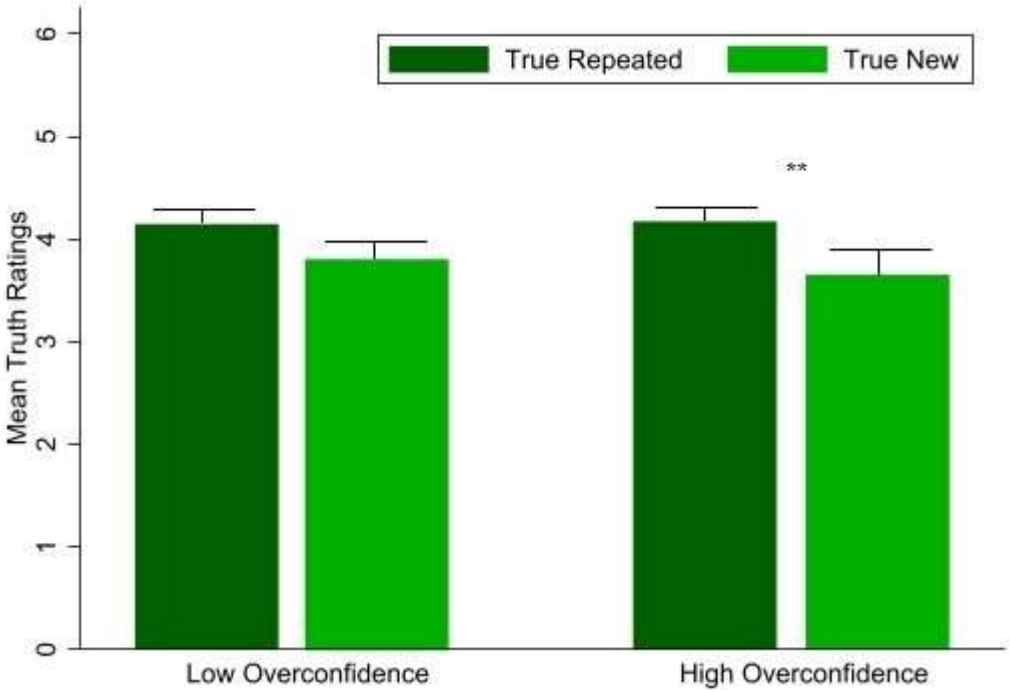
Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: True Repeated = 4.47 (SE = 0.1); True New = 4.32 (SE = 0.1). For high overconfidence subjects: True Repeated = 4.52 (SE = 0.1); True New = 4.28 (SE = 0.11). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.3 - Mean Truth Ratings for False Medical Statements by Overconfidence in Medicine



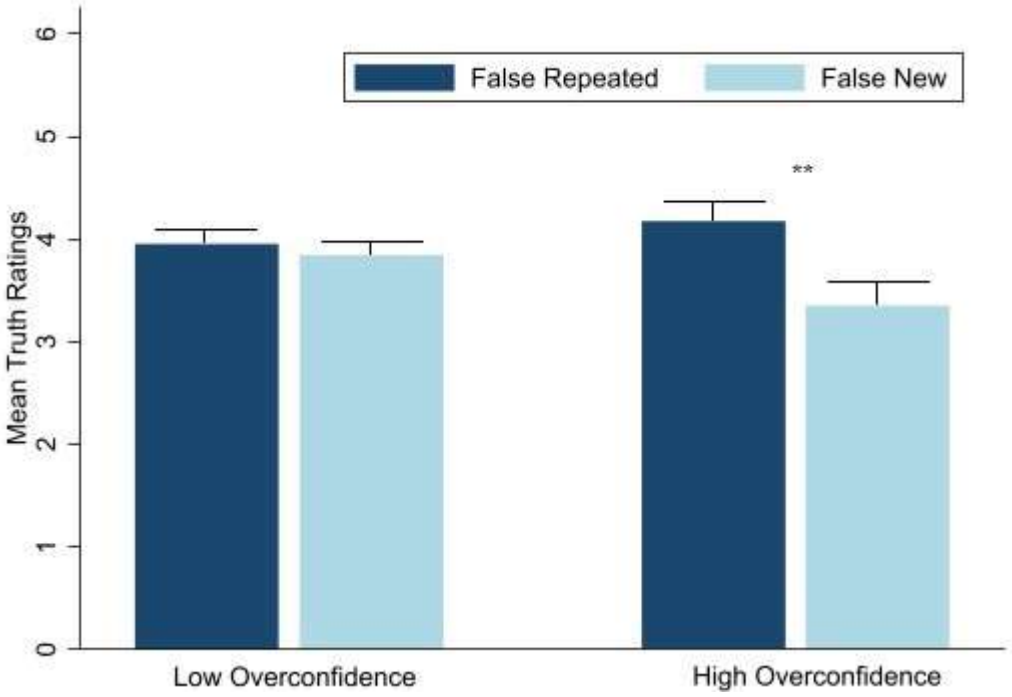
Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: False Repeated = 3.48 (SE = 0.14); False New = 3.07 (SE = 0.16). For high overconfidence subjects: False Repeated = 3.49 (SE = 0.15); False New = 3.08 (SE = 0.15). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.4 - Mean Truth Ratings for True Medical Statements by Overconfidence in Medicine



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: True Repeated = 4.15 (SE = 0.17); True New = 3.81 (SE = 0.2). For high overconfidence subjects: True Repeated = 4.17 (SE = 0.18); True New = 3.65 (SE = 0.26). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.5 - Mean Truth Ratings for False Economic Statements by Overconfidence in Economics



Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: False Repeated = 3.95 (SE = 0.17); False New = 3.84 (SE = 0.19). For high overconfidence subjects: False Repeated = 4.17 (SE = 0.2); False New = 3.35 (SE = 0.25). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.6 - Mean Truth Ratings for True Economic Statements by Overconfidence in Economics

Note: The bar chart is based on the data gathered in the experiment. The exact mean truth ratings are as follows: For low Overconfidence subjects: True Repeated = 4.98 (SE = 0.13); True New = 5.16 (SE = 0.11). For high overconfidence subjects: True Repeated = 5.09 (SE = 0.15); True New = 5.20 (SE = 0.13). Tail length on top of bars represents the standard error. Statistical significance of difference between means is marked by asterisks above bars (* = 10%; ** = 5%; *** = 1%).

II.7 – Linear Regression Formulas for Relationship Between Familiarity and Time Spent on Truth Ratings

The following 4 formulas describe the linear regressions performed to further examine the relationship between being highly familiar with a field and the time spent on rating the truth status of false and true statements from said field.

1. Relationship between high familiarity in medicine and time spent on truth rating of false medical statements:

$$TimeMedFalse_i = \alpha \cdot MedH_i + \beta \cdot Age_i + \gamma \cdot Gender_i + \delta \cdot English_i + \epsilon$$

2. Relationship between high familiarity in medicine and time spent on truth rating of true medical statements:

$$TimeMedTrue_i = \alpha \cdot MedH_i + \beta \cdot Age_i + \gamma \cdot Gender_i + \delta \cdot English_i + \epsilon$$

3. Relationship between high familiarity in economics and time spent on truth rating of false economic statements:

$$TimeEconFalse_i = \alpha \cdot EconH_i + \beta \cdot Age_i + \gamma \cdot Gender_i + \delta \cdot English_i + \epsilon$$

4. Relationship between high familiarity in economics and time spent on truth rating of true economic statements:

$$TimeEconTrue_i = \alpha \cdot EconH_i + \beta \cdot Age_i + \gamma \cdot Gender_i + \delta \cdot English_i + \epsilon_i$$

The four variables $TimeMedFalse$, $TimeMedTrue$, $TimeEconFalse$, and $TimeEconTrue$ stand for the time spent by each subject i on the truth rating of medical, false and true, and economic, false and true statements respectively. $MedH$ and $EconH$ are dummy variables taking a value of 1 if a subject i is highly familiar with medicine or economics respectively. Age is a continuous variable for the age of subject i. $Gender$ is a dummy variable for the gender of subject i, taking a value of 1 if the subject is male. $English$ is a categorical variable describing the English proficiency of subject i. ε_i is the error term for subject i.

Regression 1: *Relationship between high familiarity in medicine and time spent on truth rating of false medical statements*

VARIABLES	Time Spent on TR
High Familiarity	-0.234
	(0.956)
Age	0.261***
	(0.074)
Gender	-0.377
	(1.104)
Complete u. English	-0.621
	(1.344)
Mother Tongue English	-3.269**
	(1.383)
Constant	2.743
	(2.104)
Observations	118
R-squared	0.213

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regression 2: *Relationship between high familiarity in medicine and time spent on truth rating of true medical statements*

VARIABLES	Time Spent on TR
High Familiarity	0.769
	(2.398)
Age	0.343**
	(0.136)
Gender	-1.811
	(2.032)
Complete u. English	1.123
	(2.215)
Mother Tongue English	1.283
	(3.084)
Constant	-0.412
	(3.973)
Observations	118
R-squared	0.080

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regression 3: *Relationship between high familiarity in economics and time spent on truth rating of false economic statements*

VARIABLES	Time Spent on TR
High Familiarity	-2.822*
	(1.617)
Age	0.384**
	(0.168)
Gender	-4.297*
	(2.529)
Complete u. English	1.941
	(2.338)
Mother Tongue English	-1.289
	(1.668)
Constant	-0.464
	(4.348)
Observations	118
R-squared	0.170

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regression 4: *Relationship between high familiarity in economics and time spent on truth rating of true economic statements*

VARIABLES	Time Spent on TR
High Familiarity	-1.453
	(1.091)
Age	0.174***
	(0.059)
Gender	-3.314***
	(1.070)
Complete u. English	-0.223
	(1.694)
Mother Tongue English	-1.012
	(1.701)
Constant	4.631**
	(2.026)
Observations	118
R-squared	0.120

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1