

CHANG, YOOJIN, Ph.D. Is the Plausibility Account of the Illusion of Truth Effect Plausible? (2019)
Directed by Dr. Peter F. Delaney. 119 pp.

Repeated statements are more likely to be judged as true compared to statements that have not been repeated. This phenomenon is known as the Illusion of Truth effect. The most studied theory is that fluency induced by repetition gives an illusion that otherwise ambiguous statements are truthful. The two experiments in this dissertation tested the possibility that fluency might be supplemented by information about plausibility – that is, the presence or absence of relevant information in memory. The main dependent variable for the experiments was truth-confidence rating, which was a composite of the truth value and the confidence level for each rating reaction time. Various measures and manipulations of fluency (e.g., clarity, number of propositions, repetition) and plausibility (e.g., proposition plausibility, content valence) were included. Experiment 1 showed that despite repeated exposure of similar lexical features, contents that contradicted the target statement decreased truth-confidence ratings. Experiment 2 showed that the minimum plausibility rating of the propositions was a better predictor of veracity judgments compared to any of the reaction time measures. The results suggest that plausibility could be an important contributing factor in the Illusion of Truth effect, and possibly other related effects as well.

IS THE PLAUSIBILITY ACCOUNT OF THE ILLUSION OF TRUTH EFFECT
PLAUSIBLE?

by

Yoojin Chang

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements of the Degree
Doctor of Philosophy

Greensboro
2019

Approved by

Committee Chair

APPROVAL PAGE

This dissertation written by YOOJIN CHANG has been approved by the following committee of the Faculty of The Graduate School at the University of North Carolina at Greensboro.

Committee Chair _____

Committee Members _____

Date of Acceptance by Committee

Date of the Final Oral Examination

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	v
CHAPTER	
I. BACKGROUND AND MOTIVATION	1
Support for the Fluency Account	4
Stimulus Ambiguity and Novelty as Limitations on Tests of Alternative Accounts	8
Difficulties in Testing Fluency Accounts and Measuring Fluency	12
Proposing a Plausibility Account	15
II. THE PRESENT EXPERIMENTS	21
Pilot Study	23
Experiment 1	25
Experiment 2	34
III. GENERAL DISCUSSION	48
Assumptions of the Plausibility Account	48
Post Hoc Analysis Results	50
Implications and Future Directions	54
Conclusion	57
REFERENCES	59
APPENDIX A. TABLES AND FIGURES	66
APPENDIX B. PILOT STIMULI	85
APPENDIX C. EXPERIMENT 1 STIMULI	99
APPENDIX D. EXPERIMENT 2 STIMULI	114

LIST OF TABLES

	Page
Table 1. Anchors and Questions used for Judgment Types in all Experiments	66
Table 2. Experiment 1 t-tests for Pre-exposure Sentence Category Differences by Repetition for Confidence-Veracity Judgments (CVJ).	67
Table 3. Experiment 2 t-tests of Confidence-Veracity Judgments, Veracity Judgment Reaction Time for Once Repeated – No Repetition by Clarity and Proposition Number.	68
Table 4. Experiment 2 Correlation Coefficient Matrix.....	69
Table 5. Regression Coefficients for Reaction Time Data.	70
Table 6. Regression Coefficients for Plausibility Data.....	71
Table 7. Regression Coefficients for Plausibility and Reaction Time Data.	72

LIST OF FIGURES

	Page
Figure 1. Diagram of Experiment 1 Procedure.	73
Figure 2. Experiment 1 ANOVA Results of Confidence-Veracity Judgments by Repetition for Pre-exposed Sentences.	74
Figure 3. Histogram of Reaction Time Frequency for Experiment 1 Veracity Judgment Reaction Time.....	75
Figure 4. Histogram of Reaction Time Frequency for Experiment 1 Confidence Judgment Reaction Time.....	76
Figure 5. Experiment 1 ANOVA Results of Veracity Judgment Reaction Time by Repetition for Pre-exposed Sentences	77
Figure 6. Experiment 1 ANOVA Results of Confidence Judgment Reaction Time by Repetition for Pre-exposed Sentences	78
Figure 7. Diagram of Experiment 2 Procedure.....	79
Figure 8. Experiment 2 ANOVA Results of Veracity Confidence Judgments by Repetition and Clarity for Statements with 2 and 3 Propositions.....	80
Figure 9. Histogram of Reaction Time Frequency for Experiment 2 Veracity Judgment Reaction Time.....	81
Figure 10. Histogram of Reaction Time Frequency for Experiment 2 Confidence Judgment Reaction Time.....	82
Figure 11. Experiment 2 ANOVA Results of Veracity Judgment Reaction Time by Repetition and Clarity for Statements with 2 and 3 Propositions.....	83
Figure 12. Experiment 2 ANOVA results of Confidence Judgment Reaction Time by Repetition and Clarity for Statements with 2 and 3 Propositions.....	84

CHAPTER I

BACKGROUND AND MOTIVATION

Imagine that your friend tells you that he thinks that Elvis Pressley is still alive and living in Greensboro. How is it that we are so quickly able to reject this claim, without any apparent retrieval of other facts? It could be true, but it's probably not – and we can assess that fact very quickly on the basis of our other knowledge. Similarly, if I say that coffee is a popular drink in Trinidad, you may find this very plausible and likely to be true even if you lack any first-hand knowledge about coffee in Trinidad. How do we make these kinds of quick decisions about what is true and false when we don't know for sure?

One answer is that we use information about how easily we can process a claim to decide its truthfulness. This theory arose to explain Hasher, Goldstein, and Toppino's (1977) *illusion of truth* effect with ambiguous statements (i.e., statements where the actual veracity was unknown). They found that repeated statements were more likely to be judged as true compared to statements that were only presented during the second session (i.e., new), regardless of the actual statement veracity. In other words, ambiguous information was mistaken for true when it was repeated. The illusion of truth effect has been replicated many times since then, and the dominant view has been that repetition makes the item easier to process (*fluent*) than ambiguous statements that were not

relationship between the item and other information in a person's experience that controls decisions about truth for ambiguous items.

Besides one very recent development (specifically, the referential theory of Unkelbach & Rom, 2017), the currently dominant view of the cause of illusion of truth is based on fluency. Fluency theories propose that as statements are repeated, they become easier to process, and consequently are judged as more likely to be true (e.g., Alter & Oppenheimer, 2009; Koch & Forgas, 2012; Killer, Lloyd, & Westerman, 2008; Shapiro, 1999; Scholl, Griefender, & Bless, 2014; Sunbar, Kardes, & Wright, 2015; Unkelbach, 2007; Unkelbach & Stahl, 2009; Unkelbach, Bayer, Alves, Koch, & Stahl, 2011). Consistent with this view, other research shows that the exact statements need not be repeated, and merely exposing the participants to a relevant information can induce the same effect (e.g., Arkes et al., 1991; Ozubko & Fugelsang, 2011). For example, when shown information related to China that was irrelevant to the target statement, participants were likely to judge the target statement to be more truthful. Furthermore, simply making statements more legible induces inflated truth judgments (e.g., Reber & Schwarz, 1999). Participants were more likely to judge a statement as true when the font size was bigger, when the stimuli rhyme (e.g., McGlone & Tofigbakhsh, 2000), and when the speaker was more confident and articulate when describing the same information as opposed to mumbling it (Brennan & Williams, 1995; Zürn, & Topolinski, 2017).

However, the illusion of truth is a very large, robust effect, and there may be additional mechanisms that contribute to the effect. In this project, I tested the plausibility of an additional mechanism: one based on plausibility. Specifically, I suggest

that when we encounter a statement and need to judge whether or not it is true, we do a fast comparison with information that is readily available in memory, and to the degree that it is consistent, we judge statements to be more likely to be true. Plausibility produces the illusion of truth because the best match for a statement is often the same information, particularly if it was seen recently and frequently. However, other relevant information might also affect our truth judgments. I do not intend to argue that fluency plays no role in illusion of truth; however, I think plausibility may sometimes provide as much or more information than fluency.

One can wonder why illusion of truth research was not directed along this path earlier. Is plausibility another term to describe what researchers have been calling conceptual fluency (i.e., the content of the item is easily processed)? One possibility is that using related information is often slow, because it takes time to retrieve relevant information and logically weight it to use as a basis of judgment. Indeed, central route persuasion is often a slow process of reasoning through the available evidence about a claim (e.g., Kahnemann, 2011; Petty & Cacioppo, 1981), entailing systematic retrieval and evaluation of concepts. In contrast to this slow, deliberative approach, I proposed that the speed to make a plausibility judgment can sometimes be quick and intuitive. Later, I will argue that earlier plausibility judgment work by Reder (1982) with sentence verification supports the possibility that we can make quick, intuitive truth judgments based on a very rapid plausibility evaluation, and then use it as basis for making such judgments.

Thus, the two main variables necessary to dissociate fluency and the plausibility account are speed and content—speed being the time it takes for one to process an item and content being the relevant information that is highly available in memory. I propose that the content of the statement is important: people will need to pool relevant information that they know of to make judgment based on plausibility. Furthermore, I propose that such plausibility judgments are made quickly and are a primary source of truth judgments. However, this is not what the fluency account would predict: content is less important as long as the statement is easily processed. In the following sections of the paper, I will review the evidence for the fluency account in more detail, and then argue that our use of ambiguous, novel statements has prevented us from testing alternative accounts to the fluency account and proposed to test the relative value of fluency and plausibility accounts.

Support for the Fluency Account

The fluency account has been very successful in providing explanations for phenomena in illusion of truth (e.g., Dechêne et al., 2010). It assumes that people do not scrutinize the statements too deeply, but rather make a kind of snap judgment based on how easily the statement is processed. One piece of evidence that repetition can be driven by fluency or familiarity is that the effects of repetition on judgments are not just limited to belief judgments (e.g., Bruett & Leynes, 2015; Garcia-Marquez, Prada, & Mackie, 2016; Kurilla & Westerman, 2008). Similar effects are observed in other types of judgments. For instance, when an item is repeatedly presented, the item is more preferred to in comparison to new items (e.g., Bornstein, 1989; Foster, Fabi, & Leder,

2015; Moreland & Topolinski, 2010; Zajonc, 1968) and repeatedly presented problems are more likely to induce higher judgments of insight (e.g, Topolinski & Strack, 2009; Topolinski & Reber, 2010).

More broadly, fluency is part of a family of effects that occur with little conscious thinking. As Kahnemann (2011) described, there are often two modes of thinking: System 1 and System 2. System 1 is an automatic system that operates with little to no effort quickly which means it lacks any voluntary control on our part, whereas System 2 is an effortful system that needs concentration, attention allocation, and effortful processing to allow for complex thinking. So, in the illusion of truth paradigm, those who judged the statements to be true after repetition even though they are false are engaging in a System 1 processing. The statement is no longer novel but familiar, so without scrutinizing the statements' validity – System 2 would do the scrutinizing – System 1 would help the participants conclude that the statement is true.

Consistent with a role of System 1 in making truth judgments, when conscious resources are reduced the illusion of truth increases. For example, when participants are taxed with an attention-dividing task (e.g., Skurnick, Yoon, Park, & Schwartz, 2005) or when the participants lack motivation for further thinking (e.g, Garcia-Marques, Silva, & Mello, 2016) they are more likely to judge repeated statements as true compared to those who had full mental capacity and motivation during judgment. So, it is not hard to agree that the “illusory” part of the illusion of truth phenomenon is that participants are not fully scrutinizing the validity judgment of the statements. Instead, the illusory sense of familiarity induced by repetition is causing the truth judgment inflation.

Before the fluency account, people explained the illusion of truth in terms of item familiarity and its dissociation from the source of the items (e.g., Bacon, 1979; Begg, Anas, & Farinacci, 1992; Henkel & Mattson, 2011; Parks & Toth, 2006). They assumed that there is a sense of familiarity that repetition brings to the statement which separates the repeated statement from the novel statement (e.g., Vosoughi, Roy, & Aral, 2018), and the sense of familiarity is not tied to a valid reason, so it can easily be attributed to anything: in this case to being a true statement (e.g., Kelley & Lindsay, 1993; Kelley & Rhodes, 2002; Scholl, Griefeneder, & Bless, 2014; Wang, Brashier, Wing, Marsh, & Cabeza, 2016). For example, when the participants see a statement repeatedly along with an indication about its actual veracity (i.e., if it is true or false), they are more likely to remember the veracity value along with the statement, allowing them to give a correct truth judgment even after a week (e.g., Brown & Nix, 1996). However, when asked to rate the veracity after three months, the veracity value (i.e., that the statement is false) was lost, but the feeling of familiarity towards the statement remained, leading the participant to judge the statement to be true.

The field later began to refer to these familiarity effects as arising from fluency. What Dechêne and colleagues (2010) call fluency is what Kahnemann (2011) would call cognitive ease. When there is less effort required or any need for strained attention and control, it is easy to process (i.e., fluent). When it requires elaborate thinking as it requires systematic cognition, it is hard to process (i.e., disfluent). The fluency account suggests that items that are repeated are more easily processed (i.e., fluent) and are therefore judged as more truthful than those that are not repeated (e.g., Dechêne, Stahl,

Hansen, & Wänke, 2009). The fluency account proposed that a fluent statement is associated with truth and therefore more likely true and a disfluent statement is associated with the untruth and therefore judged as false (Schwarz, 2004; Alter & Oppenheimer, 2009; Killer, Lloyd, & Westerman, 2008; Wänke, & Hansen, 2015; Westerman, Lenska, & Olds, 2015; Whittlesea, 2002). Their idea is that fluent material must mean there is ample frequency of past exposure to the idea. When people lack memory regarding the sources of the repetitions, it may be taken to imply a social consensus (Alter & Oppenheimer, 2009; Schwarz, Sanna, Skurnick, & Yoon, 2007; Bacon, 1979), and social consensus means a higher probability of the material being true.

A more recent development in the fluency theory emphasizes that there must be relative/unexpected fluency differences between the repeated statements and the new statements to observe the effect (e.g., Dechêne et al., 2009, 2010; Hansen, Dechêne, & Wänke, 2008; Jiang & Hong, 2014; Ozubko & Joordens, 2007). This is especially prominent in a within-list design, where there are repeated statements and new statements in the same list, as opposed to a between-list design. The repeated statements are at an advantage as they are relatively more fluent than statements that are new. As the participants are unaware of why they are more fluent (i.e., they do not recall that they were repeated), the participants infer that the fluent statements must be true.

In conclusion, the fluency account is now well-established and explained in ways that are similar to other cognitive phenomena, such as the mere exposure effect (e.g., Zajonc, 1968) and the false fame effect (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989). There have been few serious challenges to the fluency account (e.g., Unkelbach & Rom,

2017), and it does a good job of explaining why the judgments can be made quickly, and why they are sensitive to repetition.

Stimulus Ambiguity and Novelty as Limitations on Tests of Alternative Accounts

In this section, I will argue that content is an important piece for understanding the illusion of truth effect, but the way previous researchers have studied illusion of truth prevented observing such effects. The plausibility account assumes that we can quickly compare novel information to highly available and relevant knowledge, and that when there are many repetitions of a given statement, it tends to be the most highly available information. Thus, I will argue that the use of only ambiguous statements and mostly exact repetitions rather than providing relevant information made it impossible to distinguish the contributions of fluency and plausibility.

In the plausibility account, when asked to make a truth judgment about a statement, what one knows about the statement primarily determines the answer as shown in the meta-analysis. If she knows that the earth is round, no matter how many times “the earth is flat” is presented, the answer will always be false because a vast number of earlier contradictory statements exist in our memories that are readily available. However, when there is no relevant knowledge about whether the statement is true or false, repeating it (e.g., Bacon, 1979; Hasher et al., 1977; Garcia-Marquez et al., 2016; Moritz et al., 2012) or reading relevant information (e.g., Arkes et al., 1991) makes the statement itself the most available information about it, so it seems true. For example, when she is not sure how much percentage of the air is nitrogen but does know that there is nitrogen in the air, then when the statement “nitrogen has a volume of 48%” is

repeatedly shown, she is likely to judge it as true even though it is wrong (the correct answer is 78%). The judgment is likely to depend on what she knows about the subject when she does not have a definitive answer of right or wrong. It is likely that such judgments are based on how plausible the statement is given what she already knows about the subject matter. In fact, that was how illusion of truth researchers explicitly excluded information for which we already knew the answer, because then people could just retrieve that the statement was true or false (e.g., Hasher et al., 1977; Bacon, 1979). And the meta-analysis found that the actual veracity does not affect the subsequent truth judgment as much as the repetition of the statements if the statements were ambiguous in their veracity (Dechêne et al., 2010). This suggests that participants can determine the obviously true and false statements, since repetition of the statements only affected participants when the statements could be judged as either true or false.

However, there are some studies that show blatantly false statements (e.g., Lead is lighter than Aluminum) can sometimes be judged as true even when people have the knowledge to retrieve the correct answer when asked to do so. For example, Fazio and colleagues (2015) found that people will erroneously judge a false statement (e.g., Sari is the short-pleated skirt worn by Scottish men) to be true with repetition even if they could correctly answer the question (e.g., kilt). Even though people have the knowledge of the statements, there can be a partial and “good enough” match—possibly even driven by a sense of fluency—that leads participants to incorrectly accept the statements as true. Similarly, in the Moses-illusion effect, people are likely to answer “two” when asked “How many animals did Moses bring into the ark?” even though it was Noah (not Moses)

who built an ark (Erickson & Mattson, 1981). The partial match of the Moses being a biblical figure such as Noah is what likely drives the erroneous response.

Expertise may also affect illusion of truth, because experts may have enough available content knowledge to avoid the partial match. Fazio and colleagues (2015) said, “We expected that participants would draw on their knowledge, regardless of fluency, if statements contained implausible errors (pp. 1000).” A critical point is that plausibility is a subjective and idiosyncratic matter. Psychology majors who had taken at least two psychology courses still showed an illusion of truth effect for false claims about psychology (Boehm, 1994). For those who have taken an introductory class and maybe another core course, they knew “just enough to be dangerous.” However, according to Fazio et al.’s logic, psychology professors should likely show no illusion of truth effect for false psychology statements, because they are very implausible given their extensive relevant knowledge. Instead of knocking on every professor’s door to test this, I proposed to use fluent but varying content valence as statement manipulation for one of the studies (Experiment 1). For example, “Ivan Pavlov studied behavior conditioning through the saliva collected from a dog” is used as a target sentence. I could use sentences that have some features of the statement that are repeated and therefore fluent in processing but differ in terms of its support for the target statement (e.g., dogs salivate when hungry vs. there is an unpleasant odor found in dogs’ saliva).

I proposed that the content knowledge of the statements is an important piece in understanding truth judgments. Whatever is plausible to the participant is likely to be judged as true. For instance, in a study using participants with high schizotypy and those

without, participants with low schizotypy were mostly unaffected by repeating paranoid statements (e.g., The Russian government spies on their citizens through multiple surveillance cameras) compared to the participants with high schizotypy (Moritz et al., 2012). To the participants with high schizotypy, the paranoid statement seems plausible and therefore is more likely to be judged as true.

Another reason why the illusion of truth effect is mostly studied with ambiguous statements is that there is something distinct about false information compared to true information in that there is a sense of novelty. A recent study shows that false news (i.e., information that is not factual) circulates faster and tends to spread wider than factual news (e.g., Vosoughi, et al., 2018). The researchers examined approximately 126,000 tweeted stories from 2006 to 2017. There was 85-95% agreement by six independent fact-checking organizations when classifying the information as either true or false. The authors investigated how differentially the information spread among Twitter users. They found that false information spread faster, more broadly, more elaborately, and farther than true information. This effect occurred across all categories, including politics, terrorism, natural disasters, science, urban legends, and financial information. The researchers also found that false information was determined to be more novel than true information, which suggests that people¹ tend to share novel information. Whatever is a slight variation of what I know to be plausible is attributed as novel and often, these slight variations tend to be false.

¹ According to Vosoughi and colleagues (2018), both true and false news were more likely spread by robots than people as opposed to what is popularly believed.

In sum, though it has not been the focus of the research in the field, I proposed that content of the statements and having relevant content knowledge are crucial to making rapid truth judgments. I will study the repetition effect when the valence of the content is different (see Experiment 1). I also proposed that even though a statement is read easily when there is a piece of information that is a slight variation in what they already know to be plausible, participants will detect it and judge the statement to be false (see Experiment 2).

Difficulties in Testing Fluency Accounts and Measuring Fluency

This section challenges and discusses the limits to the method of studying how an increase in processing ease mediates the illusion of truth effect. My concern is that fluency cannot be easily measured and was usually not measured in illusion of truth studies. Instead, fluency was the manipulation (i.e., repetition or ease of reading). I will suggest using speed of processing during the judgment as a possible way of measuring fluency in illusion of truth effect paradigms.

One way to manipulate fluency is via the availability heuristic, which states that sometimes the frequency of available examples (e.g., Kahnemman & Tversky, 1974) or how easily an example comes to mind (Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka, & Simons, 1991) governs the judgment. For instance, Schwarz and colleagues (1991) devised a study where the main question was to identify a person's judgment about a trait category based on the frequency of instances related to the category. The participants would recall instances that they have behaved assertively and later would rate how assertive they are. One group of participants were asked to list six instances when

they were assertive, and the other group of participants were asked to list twelve instances. When the participants were required to answer twelve, they struggled to recall instances after a certain number. While the participants that were required to answer six, they could recall instances easily compared to the group that were required to answer twelve instances. After recalling the instances, both groups rated how assertive they thought they were. The participants in the twelve instances group judged themselves to be less assertive than the participants in the six instances group because they had a difficult time coming up with instances of assertiveness to meet the requirement: twelve instances. When asked to come up with instances when they were not assertive, the participants in the twelve instances group thought they were more assertive than the participants in the six instances group. The basis for the participants' judgment on their assertiveness did not depend on the actual number of instances recalled but whether they felt it was easier or harder to come up with the examples. So, it can be established that fluency is the subjective feeling of ease.

But how would one measure fluency when judging a statement's truth besides participant self-report? When looking into the illusion of truth effect, fluency is not a dependent variable but rather an independent variable (e.g., Alter & Oppenheimer, 2009). Traditionally, statement repetition was the main source of fluency manipulation (e.g., Difonzo, Beckstead, Stupak, & Walders, 2016; Garcia-Marquez, Silva, Rever, & Unkelbach, 2015; Hasher et al., 1977; Johar & Rogeeven, 2007; Koch & Zerback, 2013; Moons, Mackie, & Garcia-Marquez, 2009). However, research within the area found that the illusion of truth effect could be found when there is no exact repetition of the

statements. In those studies, fluency is manipulated so statements could be visually easy to read (Reber & Schwarz, 1999), linguistically proficient (Brennan & Williams, 1995), or rhyming (McGlone & Tofiqbakhsh, 2000). Fluency was not determined by the participants, but rather by the experimenters' manipulation.

Though fluency is assumed to be the main factor for illusion of truth effects, the literature is unclear about the relationship between the judgment latency and judgment veracity. The meta-analysis discusses stimulus presentation time, delay between first judgment and last judgment, and interval times between stimuli presentation, but nothing about the time to make the judgment itself (Dechêne et al., 2010). In their paper demonstrating retrieval fluency (i.e., how easily something is retrieved) as the basis for confidence in answers to knowledge-based questions, Kelley and Lindsay (1993) showed that there was a negative correlation between response latency (i.e., how long it took for the participants to answer) and confidence (i.e., how confident they are that the answer they gave is correct). The confidence ratings were low when the participants took longer to respond. They argue that "the ease and speed with which supporting evidence comes to mind may also contribute to confidence (pp. 19)." Seeing confidence is also a type of meta-cognitive judgment (e.g., Alter & Oppenheimer, 2009), Kelley and Lindsay claimed that the fluency (i.e., ease and speed) of evidence generation can be used as basis for making a confidence judgment, I don't see why that cannot be incorporated for truth judgments.

If fluent statements seem fluent because they are processed quickly, and if fluency is what makes for an illusory truth judgment, then logically disfluent statements should

be judged slowly – there shouldn't be fast false judgments. However, Unkelbach and Rom (2017) found that participants can quickly discern that a presented statement differs in only a single meaningful way from an earlier-presented statement. Such statements should be processed quickly because they are highly familiar, and hence fluent, yet they are still judged as false very rapidly. This begs the question as to whether fluency predicts that the whole statement should be processed quickly, or if each of the parts in a statement should be processed quickly. I intend to investigate that (see Experiment 2) by using speed as a dependent variable.

Proposing a Plausibility Account

The main goal of this dissertation is to propose mechanisms other than fluency that could explain illusion of truth effects—perhaps not to replace them altogether, but to complement them. Specifically, the plausibility account proposed that when confronted with a statement whose veracity needs to be examined, we rapidly compare the statement to readily available information in memory. To the degree that the most readily available information is consistent, we tend to think it is true. To the degree that the most available information is inconsistent, we become skeptical. The plausibility account suggests that repeated statements are judged as more truthful because repetition of the statements increases the availability of relevant information (i.e., the same statement), and hence the sense of plausibility. The reason why repeated statements and conceptually relevant information feels fluent may be that the information is plausible and therefore believable, and that should only increase with more repetition of strongly relevance information.

Thus, fluency could even be the result of plausibility, albeit in the special case where other relevant information is mostly unavailable.

The remainder of the document proposed experiments that will get at the plausibility of a plausibility account. Overall, I concur with the existing literature in that the illusory truth judgments are made when there is a lack of veracity information (e.g., Brown & Nix, 1996) and the judgments are made because the statements are familiar (e.g., Bacon, 1979). I can see how repeated statements are perceived to be fluent and that it may lead to an illusory truth judgment (e.g., Alter & Oppenheimer, 2009). Fluency accounts can explain how perceptually easy to read stimuli are more likely to be judged as true, however that effect size is on average small ($d = .1$) compared to the overall illusion of truth effect size ($d = .3$) (Dechêne et al., 2010). This suggests that there is residual variance that needs to be explained.

Thus, I proposed content and speed can be important factors that together can help elucidate the remaining effect. Recently, Unkelbach and Rom (2017) found that new statements were 640 ms faster compared old statements. As mentioned in their paper, though slow judgments do not necessarily mean it is disfluent, statistically, fluent statements tend to be judged faster (e.g., Whittlesea, 1992). Unkelbach and Rom also found that statements that were judged true were more likely to be judged in a faster manner than statements that were judged as false which is consistent with what fluency account would predict. The critical piece to their paper was their manipulation of old-contradicting statements. For example, when participants studied statements such as “The world’s most poisonous snake is the Australian Coastal Taipan,” but was shown

Australian Inland Taipan, participants responded relatively faster than the new statements but was able to correctly identify as false when asked to judge true or false. This would suggest that participants easily processed the statements due to repetition but was able to correctly and quickly find an error.

Not only is studying the content of the statement crucial, what you do with the content influences truth judgments. In their fourth experiment, Unkelbach and Rom (2017) showed statements such as “Most accidents occur close to weekends” as the target statement that required truth judgments but had different study conditions for half of the participants. Half of them saw a statement that implied the target statement such as “Most accidents occur on Mondays.” The results showed that disregarding the validity of the studied statement, it resulted in inflated truth judgments because it was implied.

Not unlike the implied statements affecting truth judgments, I proposed that participants can use their knowledge base and what is learned from the content of the statement to connect the dots to make a plausible inference. The plausibility account of the illusion of truth effect was inspired by research on “sentence verification” by Reder (1982). Sentence verification occurs when one has read relevant information and must now indicate whether a given statement is consistent or inconsistent with what was read. This is strikingly close to making a truth judgment, except that relevant information was provided by experimenters in the case of sentence verification, but the only relevant information was the statement itself in the case of the illusion of truth studies.

In contrast to sentence verification, sentence recognition is a probe match that requires specific retrieval of the target. Reder (1982) argued that sentences can be

verified by one of two ways: direct retrieval (e.g., direct match with the same target in memory) or by computing plausibility. She hypothesized that fact retrieval is less efficient sometimes than comparing plausibility (or inferring) especially at a delay. Plausibility is more effective in that direct retrieval needs a specific match (e.g., facts about Hope diamond) that might be inaccessible at a given moment. Plausibility judgments can be made using whatever is easily accessed at a given moment. She hypothesized and found (Experiment 2: Reder, 1982) that direct retrieval is not always the preferred/default strategy, nor is it effectively preferred over plausibility judgments. Her logic was that plausibility judgments will be preferred when there is no verbatim match found between the sentence probe and what is stored in memory.

In her studies, Reder (1982) used two types of judgments: plausibility and direct retrieval. She used a sentence recognition task, which asked participants to make judgments of whether they recognize the sentence or not (i.e., direct retrieval). She also used a sentence verification task, which asked participants to make judgments of whether the sentence is sensible given the background story they have studied with the (inferred) sentence (i.e., plausibility). Participants made judgments of whether they recognized sentences or thought the sentences were plausible after having read stories of moderate interest. The sentences were either what was already stated in the story or sentences that were not. The statements differed in the degree of plausibility: highly plausible, moderately plausible, and implausible. She commented that components of the sentences were pulled from the story that participants originally studied but was all together implausible. This was carefully designed to avoid participants making plausibility

judgments out of lexical unfamiliarity. The judgement types (i.e., recognition or plausibility) and when the judgment occurred differed among participants. A group of participants made the judgments (either recognition or plausibility) immediately after the study session, another group of the participants made the judgments (either recognition or plausibility) 20 minutes after the study session, and the other group of participants made the judgments after a two-day delay.

Though there are more results² than that which contribute to understanding plausibility, the results that are most relevant to the plausibility account of illusion of truth are the reaction time results. After a delay, participants' speed for making plausibility judgments were much faster compared to participants' speed for making recognition judgments. Reder (1982) argued that participants were initially faster and more accurate in the immediate condition because intact verbatim traces of the story were available. However, after a delay, this trace weakened and therefore it took longer to make accurate judgments. Plausibility judgments do not require such a verbatim match as long as some trace of the story was represented and accessed.

One can see the parallel between the sentence verification task and illusion of truth tasks with Reder's paradigm. Participants were asked to make judgments based on memory and the inferences made from what is remembered (Reder, 1982). When the memory trace is weakened, participants and are constricted in response time, they were

² With more time delay, participants' difference between recognition and plausibility increased. In that, participants in the recognition judgment group were less accurate after a delay than participants in the plausibility judgment group. Not only did patterns of accuracy responses change after a delay but so did reaction measures for both group of participants.

more likely to make a plausibility judgment faster and more accurate compared to when having to retrieve. Illusion of truth studies demonstrate a similar process whereby participants will make the truth judgment based on what is plausible when the trace memory is weakened — and they can do it quickly.

Reder's (1982) claim that a plausibility mechanism could explain how people rapidly made judgments in sentence verification could be extended to illusion of truth experiments. Therefore, when there are more pieces to be used as cue to pool from already stored knowledge, the more likely the statement will be judged as true. As long as there is a single proposition that is implausible, participants will be able to make a "false" judgment based on it.

CHAPTER II

THE PRESENT EXPERIMENTS

While my future research will develop more specific mechanisms underlying plausibility, for the dissertation project, I compared the relative success of fluency and plausibility mechanisms more generally under circumstances likely to elicit plausibility effects. I specifically examined whether or not the speed to make the judgments and availability of relevant content knowledge about the statements influenced the final truth judgments. The series of experiments were designed to determine whether a plausibility mechanism is needed at all, or whether fluency³ alone can handle all of the results.

There were two main experiments to test the hypotheses and one pilot experiment that was used to norm the material statements. The two main experiments consisted of a study (exposure) phase and a rating phase. During the study (exposure) phase, participants read the statements at their own pace. During the rating phase, participants either read a new statement or a version of the old statement or made truth judgments along with confidence judgments. The dependent variables for the two main experiments were the proportion of truth judgments (true or false), the confidence ratings about their previous truth judgment (1 – 5), and the reaction time during the rating (See Table 1 for detailed scale information). Though early studies used a long interval between the study and rating phases (e.g., Bacon, 1979; Hasher et al., 1977), recent studies found that the

³ I both manipulated and measured fluency in all experiments.

interval themselves do not mediate the illusion of truth effect (e.g., Nadarevic & Erdfelder, 2014) and that effect could be found without a delay (e.g., Ozubko & Fugelsang, 2011). Therefore, for the experiments, there was no delay between the study phase and the rating phase.

Experiment 1 focused on the supporting value of pre-exposed information about a statement by providing two associated sentences that contained either supporting (positive), contradicting (negative), or non-informative (neutral) statements related to the target statement. After studying these pre-exposed informative sentences, participants made judgments about the truth of the target statements. The fluency account implies that the valence of the sentences does not matter as long as they are repeated; any presented information that speeds future processing should lead to fast judgments, and hence seem true. In contrast, the plausibility account directly predicts that the content valence of readily-available earlier claims would be an important factor in truth judgments, such that positive and possibly neutral information would increase illusion of truth, while negative information should decrease it.

Experiment 2 measured plausibility of the parts of sentences directly. Participants studied a list of statements during the initial study phase. During the subsequent rating phase, participants rated the validity of a version of the statement that differed in clarity (e.g., awkward or identical to target) and the plausibility of the propositions assigned to that target statement. The fluency account directly predicts that the driving force behind inflated truth judgments should be factors that affect speed of processing, such as the clarity of the statements, length of the statements, and frequency of repetition of the

statements. In contrast, the plausibility account directly predicts that the lowest plausibility rating of the component propositions should be most important factor. Obviously, these accounts are not mutually exclusive, and so therefore they could both turn out to be true under different circumstances.

Pilot Study

The purpose of the pilot study was to norm the statements for future experiments. As previous research shows, statement ambiguity is a crucial factor in producing the illusion of truth phenomenon (e.g., Dechêne et al., 2010). When a statement is obviously true or false, there is usually little to no illusion of truth effect found. Therefore, it is important to use statements that are of an ambiguous nature (i.e., participants cannot easily tell if they are true or false).

Method

Participants. I planned to run 18-20 participants per group based on previous norming studies. Fifty-seven students from UNCG participant pool were recruited. Three participants who did not complete the task were excluded from the data. There were 18 participants in each group.

Materials and Design. The materials consisted of 96 trivia statements including two versions of incorrect statements making a total of 282 statements. The statements were collected from areas such as history, art, science, literature, film, music, social issues, and math from a distractor task used in a separate project with a collaborator, Peter Verkoeijen and his lab. An example of the trivia statements is, “Mel Blanc voiced the cartoon character Bugs Bunny.” The two versions of incorrect statements are the

same sentence with one element switched to an alternative. For instance, in the example of above, the incorrect statements mentioned Don Messick or Joseph Barbera instead of Mel Blanc. The statements are provided in Appendix B.

I constructed three counterbalanced versions of the 96-item statement lists, each containing 32 true statements and 64 false statements. Participants were randomly assigned to one of the three lists. The order in which the statements were presented was randomized for each participant. Thus, everyone rated each statement once in either its true or one of its false versions.

Procedure. Participants were told that they would see a series of trivia statements. They were informed that some of the statements are true and some of them are false, without the exact ratio information being given. Each statement was presented with a True or False question and participants answered by clicking on the answer that best described their judgment. Afterwards, they rated how confident they were about the previous rating and the plausibility of the statements. See Table 1 for examples. The response rates for each truth judgment, confidence judgment, and plausibility judgment were recorded and were self-paced.

Results

The goal of the pilot was to norm the statements for future experiments. See Appendix B for the results and statements. I collected the accuracy rate of each statement (e.g., true items rated as true vs. true items rated as false). I collected the confidence judgment ratings, the plausibility ratings for all statements including the 96 true

statements and its two false counterparts, and the reaction time was recorded and analyzed.

The mean accuracy rating was $M = 0.50$ ($SD = 0.05$), the mean confidence rating was $M = 1.85$ ($SD = 0.02$), and the mean plausibility rating was $M = 2.79$ ($SD = 0.07$). The mean reaction time from start to first response (i.e., truth ratings) was $M = 6.53$ ($SD = 0.04$).

Experiment 1

The purpose of Experiment 1 was to examine whether people rapidly use relevant information that was previously studied to produce their subsequent truth judgments about a statement. I therefore manipulated whether or not relevant content was presented before making a truth judgment about a target statement. The pre-exposed sentences were either an exact repetition (i.e., an illusion of truth replication) of the target statement (e.g., Mel Blanc voiced the cartoon character Bugs Bunny); relevant and consistent with the statement (e.g., Mel Blanc voiced Roger Rabbit before his death in 1989), relevant and inconsistent with the statement (e.g., Bugs Bunny was voiced by Greg Burson during the 1990s), or seemingly relevant but non-informative (e.g., Bugs Bunny is a character in Looney Toons first created in the 1930s) about the veracity of the statement to be judged.

This manipulation was designed to help differentiate the effects of fluency and plausibility. The fluency account predicts that pre-exposed relevant information should speed fluent processing of the statements by making the words in them more highly primed, regardless of the actual relevance of the pre-exposed statement. Consistent with this prediction, earlier studies have shown that pre-exposed statements that overlap in

content (e.g., Arkes et al., 1991) or even briefly-shown primes intended to speed reading of statements (Brown & Marsh, 2009; Kelley & Lindsay, 1993) increase the probability of rating statements as true compared to new statements.

In contrast, the plausibility account predicts that people should be sensitive to the content of the pre-exposed information. Therefore, they should be more likely to accept as true a statement that has relevant and consistent information about it in memory than a statement that has relevant but inconsistent information about it in memory.

Furthermore, this process must occur quickly if it is to be a candidate account of the illusion of truth effect.

Method

Participants. In accordance with the power analysis, 84 participants' data were used for analysis. There were 110 participants from UNCG subject pool in total but the data for the participants that were noncompliant as reflected in reaction time data were excluded. The goal number of participants was based on an a priori power analysis on G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). The analysis was conducted for an ANOVA: repeated measures, within-between interactions. The repeated measures correlation was set at $r = .02$ based on a previous study that was done in the lab on Illusion of truth. The effect size was set at $f = .25$ a medium effect based on the literature (e.g., Dechêne et al., 2010) and the power was set at $\beta = .85$. The within-subjects factor was the number of repetitions (3: 0, 1, & 2) and the between factor was the pre-exposed sentence category (4: relevant, neutral, negative, & exact). This is the lowest powered analysis that tests my most important hypothesis.

Design. The design was a 4 (pre-exposed sentence category: contradicting, neutral, confirming, exact) x 3 (number of repetitions: 0, 1, 2) mixed design. The category of pre-exposed sentences was the between-subjects factor and the number of repetitions was the within-subjects factor. The category of the pre-exposed content provided during the study phase differed among the four groups. Participants in the *exact* group saw the exact target statements. This group was a control group to replicate the illusion of truth effect. Participants in the *negative* group saw the negatively associated sentences of the target statements. Participants in the *positive* group saw the positively associated sentences of the target statements. Finally, participants in the *neutral* group saw the neutrally associated sentences of the target statements. The dependent variables were reaction time (veracity rating and confidence rating) and the mean truth-confidence judgment (ranging from -5 [very confident not true] to +5 [very confident true]).

Materials. Normed stimuli statements from the pilot study were used in this study (see Appendix B for stimulus set). Sixty statements that are near the middle plausibility point ($M = 2.61$, $SD = 0.49$) and those that participants had difficulty judging as true or false (54% accuracy, mode confidence = 1) were selected as target statements for judgments. Half of the statements were true, and the other half were false. The participants were told that there were both true and false statements but did not know the exact ratio.

Each 60 target statements each had two associated pre-exposed sentences that varied in their content valence. Native English-speaking research assistants⁴ were asked

⁴ Many thanks to Myranda Cook, Danielle Chapman, and Carson Peske for their help.

to make sentences that confirmed the target sentence (i.e., confirming), contradict it (i.e., contradicting), and were related but neither confirm nor contradict the target statement (i.e., neutral). We tried to make sentences based on true facts, but there were a few that needed to be made up.

For example, if the target sentence were, “The Hope diamond is cursed,” then an example pre-exposed content sentence in the confirming category would be, “Their owners have died upon the possession of the Hope diamond.” For the neutral category, an example would be, “The Hope diamond is blue,” and for the contradicting category, an example would be, “Diamonds are lucky in Tanzania.” Examples of pre-exposed content sentences are in Appendix C.

Procedure. During the exposure phase, the pre-exposure statements from the correct category was presented. Presentation was self-paced so the participants could advance to the next screen. Repetition of the pre-exposed sentence was manipulated in a within-subjects design varying from 0 to 2 repetitions. That is, one-third of the associated pre-exposure sentences of the target statements were shown during the exposure phase (0-repetition condition); the target statements in the 0-repetition condition were only rated during the rating phase. Another one-third were shown once during the study phase (1-repetition), and the rest were shown twice during the exposure phase (2-repetition).

During the rating phase, each of the target statements appeared with a question asking whether the statement was true or false. Then, participants rated how confident they were in the judgment. The scales were identical to those in the Pilot study (see Table

1), and the full procedure is illustrated in Figure 1. The ratings were self-paced, but the veracity judgment and confidence judgment reaction times were recorded.

Results and Discussion

My main goal was to investigate whether plausibility mechanisms can play a role in illusion of truth judgments. Plausibility predicts that the content valence of the pre-exposed sentences would affect the ratings, but fluency alone would not predict such a difference, as the lexical familiarity should increase fluency and therefore induce the effect.

Veracity-Confidence Rating⁵. I conducted a 3 (repetition: 0,1,2) x 4 (content valence: exact, confirming, neutral, contradicting) mixed ANOVA (see Figure 2 for means). The veracity-confidence ratings were calculated by multiplying the confidence ratings by their veracity ratings with the truth as a positive score and false as a negative score. The ratings therefore ranged from -5 [very confident not true] to +5 [very confident true].

As be predicted by both plausibility and fluency, there was a main effect of repetition, $F(2, 160) = 15.547, p < .001, \eta_p^2 = 0.163, MSE = 9.783$. Once repeated ($M = 1.005, SD = 1.160$) statements had significantly higher veracity confidence judgments than new statements ($M = 0.380, SD = 0.805$), $t(83) = 3.854, p < .001$, as did twice repeated ($M = 0.930, SD = 1.326$), $t(83) = 3.027, p = .003$. However, there was no

⁵ I conducted the standard illusion of truth effect paradigm with the veracity judgments only and the results were similar to that of the veracity-confidence ratings.

significant difference between the twice repeated and once repeated statements' veracity confidence judgment, $t(83) = 0.939, p = .350$.

As plausibility predicted, there was a main effect of content valence, $F(3, 80) = 18.743, p < .001, \eta_p^2 = 0.413, MSE = 1.167$. The exact sentences ($M = 1.467, SD = 0.784$) had significantly higher confidence veracity ratings than confirming sentences ($M = 0.794, SD = 0.544$), $t(40) = 3.374, p = .002$, neutral sentences ($M = 0.742, SD = 0.408$), $t(40) = 3.913, p < .001$, and contradicting sentences ($M = 0.055, SD = 0.692$), $t(40) = 6.322, p < .001$. Confirming sentences had significantly higher confidence veracity ratings than contradicting sentences, $t(40) = 3.852, p < .001$, but not for neutral sentences $t(40) = 0.353, p = .726$. Lastly, neutral sentences had higher confidence veracity ratings than contradicting sentences, $t(40) = 3.921, p < .001$. In sum, disregarding any repetition effects, exact repetitions induced higher veracity-confidence judgment ratings than any other sentences, and confirming sentences produced higher veracity-confidence ratings than contradicting sentences.

These main effects were qualified by a significant interaction of repetition and content valence, $F(6,160) = 13.766, p < .001, \eta_p^2 = 0.340, MSE = 0.629$. To follow up the interaction, the number of repetitions was separated by content valence of the pre-exposed sentences. See Table 2 for the follow up t -test results. For exact sentences, there was a significant difference between no repetition and one repetition and two repetitions. This can be considered a replication of the Illusion of Truth effect, as repetition induced an inflated truth judgment. There were no differences between one and two repetitions. For confirming sentences, there was a significant difference between no repetition and

one repetition and two repetitions, replicating the illusion of truth effect for related sentences. There were no differences between one and two repetitions. There were no repetition effects for neutral sentences, indicating merely related but non-informative material did not induce an illusion of truth effect. Finally, for contradicting sentences, only when the sentences were repeated twice were there a significant *negative* repetition effect compared to no repetition.

Averaging across all content valences, there was a repetition effect in that some repetition is better than no repetition, which both fluency and plausibility would tend to predict. However, when comparing the twice repeated and no repeated statements' veracity confidence judgments for the contradicting sentences, we see the mean ratings decreased. This is the opposite of an illusion of truth effect, in that the more contradicting statements are presented, the more likely the participants thought it to be false. As only plausibility would predict, content valence influenced the ratings.

Reaction Time. Reaction time was divided into two categories: the time it took the participants to make the veracity judgment (i.e., VJRT) and the time it took the participants to make the confidence judgments (i.e., CJRT). Previous literature has focused on the veracity judgment while asking the confidence simultaneously. However, the present studies look at the two times separately as the values indicate different concepts. The time it takes to make the veracity judgment should reflect the time it takes the participant to decide whether it is truthful or not (i.e., + or – on the current measurement scale). The time it takes to make the confidence judgments should reflect

the degree to which they are confident about the previously made veracity judgment (i.e., 1 being very not confident and 5 being very confident on the current measurement scale). The reaction time data for both veracity judgment and confidence judgments, were approximately normally distributed (see Figures 3 and 4). Therefore, the raw reaction time data were used for the analyses.

The fluency account implies that repetition decreases the time it takes to make the veracity judgment. The content valence would not make any difference as long as there were some content that was related to the target statement. In contrast, the plausibility account suggests that the more relevant the information is to the target statement, the more confident the judgments are which would manifest by speedy confidence judgment time (for review see Alter & Oppenheimer, 2009).

Veracity Judgment Reaction Time (VJRT). Fluency predicts that repetition would induce a faster VJRT. However, a 3 (repetition) x 4 (content valence) ANOVA on VJRT (see Figure 5) revealed no significant effects of VJRT for repetition, $F(2,160) = 1.201$, $p = .304$, $\eta^2 = .015$, $MSE = .047$, condition, $F(3,80) = 1.496$, $p = .222$, $\eta_p^2 = .053$, $MSE = .671$, or their interaction, $F(6,120) = 0.991$, $p = .434$, $\eta_p^2 = .036$, $MSE = .947$. Notably, judgments are made very fast (less than 2s), consistent with the plausibility account assuming a fast judgment.

Confidence Judgment Reaction Time (CJRT). Plausibility predicts that there is a difference. A 3 (repetition) x 4 (content valence) ANOVA on CJRT (see Figure 6 for means and Table 2 for t -test results) revealed a main effect of repetition, $F(2, 160) = 7.563$, $p = 0.001$, $\eta_p^2 = 0.086$, $MSE = 182.942$, where the once repeated sentences ($M =$

6.067 s, $SD = 2.217$ s) were slower than new sentences ($M = 6.407$ s, $SD = 2.091$ s). Twice repeated sentences ($M = 5.798$ s, $SD = 1.921$ s) also induced a decrease in judgment time compared to new sentences. There was no difference between once and twice repeated sentences, but the difference was in the right direction for the two repetitions being faster, $t(83) = 1.954$, $p = .054$. There was no main effect of condition, $F(3, 80) = 0.539$, $p = .657$, $\eta_p^2 = 0.020$, $MSE = 2370.372$, or interaction between repetition and condition, $F(6, 160) = 1.059$, $p = .390$, $\eta_p^2 = 0.038$, $MSE = 182.932$.

Conclusion. The Experiment 1 veracity-confidence judgment results showed that there was an effect of repetition and content valence on veracity confidence judgments. Repetition induced an increase in the veracity-confidence judgments for all content valences except for the contradicting condition, which is what was predicted by the plausibility account. As both the plausibility and fluency accounts would predict, repetition of the exact statements did increase the veracity-confidence judgments which was a replication of the immediate Illusion of Truth effect studies (e.g., Brown & Nix, 2006). What is different about this experiment compared to that of the previous literature is that the veracity judgments and confidence judgments were separated into two judgment making trials instead of combining them onto one scale. This allowed me to distinguish whether repetition affected reaction time to the veracity judgment or the confidence judgment. The reaction time data suggested that repetition influenced the confidence ratings but not the veracity judgments. Specifically, repetition seemed to decrease the time it took to make the confidence judgments about their veracity judgment

(and the veracity judgment people gave), but the time it took to make the veracity judgment was unaffected.

With Experiment 1, I have shown that plausibility might play a role in how the illusion of effect occurs. But to what degree? With Experiment 2, I proposed to understand how much plausibility would be able to account for the effect in comparison to the fluency account.

Experiment 2

The main purpose of Experiment 2 was to compare directly whether plausibility of parts of the statement predicts illusion of truth for the statement. The plausibility account suggests that people rapidly assess the plausibility of parts of a statement, so the more parts there are, the greater the likelihood that one of them seems implausible. Furthermore, it suggests that the most implausible part of the statement is the part that should drive the truth judgments. For example, if a person read a statement like, “The phrase Bob’s your uncle originated in Ireland in 1900 when Barack Obama became prime minister through nepotism,” everything likely seems plausible except that Barack Obama was in Ireland in 1900. Therefore, the whole statement seems false. For the sake of argument, the piece about US former president becoming the prime minister of Ireland was added to show plausibility’s role in the judgment. In the typical illusion of truth paradigm, the overall plausibility of the pieces is ambiguous. For ambiguous statements, the fluency hypothesis suggests that the length of the sentence and the speed of reading it matters, not the number of facts therein *per se* (although these should be correlated).

I therefore manipulated both fluency and plausibility by varying the number of propositions in a sentence and whether it was phrased in a more awkward way or an easier-to-read way. Propositions can be defined as the smallest unit of knowledge for which its veracity can be judged as either true or false (Anderson, 1974; Kintsch & van Dijk, 1978). The more propositions are in a statement, the longer it will be, making the statement harder to process. The fluency account predicts that both manipulations (i.e., number of propositions and phrasing clarity) should decrease fluency and therefore decrease the likelihood of judging the statement as true. The plausibility account would also be consistent with effects of repetition and clarity, as the match between the target at rating and study will be a deciding factor for judgment. The plausibility account predicts an effect of the number of propositions because statements that have more propositions would tend to have a higher probability of having an implausible proposition.

After making their truth and confidence judgments, people rated the plausibility of each of the propositions in the sentence. The plausibility account predicts that any implausible part of the sentence should make the whole implausible. Furthermore, the presence of an implausible part should be a better predictor than the number of propositions. By this logic, the minimum plausibility rating would be a better predictor of final veracity-confidence judgments than the time it took to make the veracity judgments as the fluency account would predict. More propositions only increase doubt because one or more of them may be implausible, not due to length *per se*. For example, “Donald Trump is a rich billionaire who was on television and became president of the United States in 2016, and owns golf courses and hotels,” is long and full of propositions, but all

of them are true/plausible, so the sentence should still be judged as true. In sum, fluency would predict the statement clarity, number of propositions, and repetition, and the time it took to make the veracity judgment would be a better predictor for final veracity-confidence judgments. However, plausibility would predict the minimum plausibility rating value and the time it took to make the plausibility rating would be the strongest predictors.

Method

Participants. Seventy-six participants' data were used for analysis in accordance with the number of participants required by the power analyses. Two power analyses were conducted. A total of 47 participants were set as goal for participant number based on an a priori power analysis on G*Power 3.1.9.2 ((Faul, Erdfelder, Lang, & Buchner, 2007) for a step-wise multiple regression with two predictors (plausibility and reaction time). A total of 76 participants were needed for an ANOVA: repeated measures, within-between interactions. The repeated measures correlation was set at $r = .02$ based on a previous study that was done in the lab on Illusion of truth. For both analyses, the effect size was set at $f = .25$ a medium effect based on the literature (e.g., Dechêne et al., 2010) and the power was set at $\beta = .85$. A total of 86 participants' data were collected, but 10 participants' data was excluded due to too fast or missing reaction time data.

Materials. The material used for study and truth judgment rating were statements that were normed in the Pilot study. See Appendix D for statement examples. The statements with low confidence ratings ($M = 1.14$, $SD = 0.45$) and accuracy rating that reflected ambiguity ($M = 0.50$, $SD = 0.16$) were collected. The goal was to select

statements that had large range of plausibility, but the majority of the statements were rated with 2's and 3's out of the 1 to 4 range ($M = 2.68$, $SD = .55$). There were a total of 36 statements, of which half were true and half were false. Of the 36 statements, 24 were used during the study phase and the remaining 12 statements were presented as foils during the rating phase.

One native English-speaking research assistant⁶ generated the awkward counterpart statements for each original statement. The instructions for the research assistants were to make the sentence lengthy, use passive voice, and use as many words as possible to make the sentence seem harder to read. For example, the sentence, “Mel Blanc voiced the cartoon character Bugs Bunny” was rewritten as, “Bugs Bunny who is a cartoon character used to be voiced by a man named Mel Blanc.” The word count was higher for the awkward statements ($M = 13.32$, $SD = 4.87$) compared to the clear target statements ($M = 10.21$, $SD = 2.70$). I evaluated appropriateness of the statements generated by the research assistants.

Another native English-speaking research assistant⁷ segmented the statements into propositions by identifying content chunks (i.e., propositions). The instructions for the research assistants were to segment the sentence with one unit of thought per segmented sentence. For example, the sentence “Mel Blanc voiced the cartoon character Bugs Bunny” was segmented to: (1) Mel Blanc is a voice actor, (2) Bugs Bunny is a cartoon character, (3) Mel Blanc voiced Bugs Bunny. The number of propositions ranged

⁶ Many thanks to Paolo Forcadela.

⁷ Many thanks to Katlynn Mabine.

from two to five. However, only statement with two and three propositions were analyzed as those were the statements that all participants saw during the experiment. I evaluated appropriateness of the statements generated by the research assistants.

Design. The main independent variables were the number of propositions within a given sentence and the clarity of the statements' phrasing. It was a repeated measures design where the clarity of the statements (direct vs. awkward), repetition (0 vs. 1) and the number of propositions (2 vs. 3) were all within-subjects factors. The main dependent variable was, as in Experiment 1, a combination of the truth and confidence judgment (-5 to +5). I also measured how long it took for the participants to make the judgments. Finally, the rated plausibility of each of the propositions that make up the target and foil sentences was also a dependent variable.

Procedure. There were two phases in this experiment: the study phase and the rating phase. See Figure 7 for an illustration of the procedure. During the study phase, participants studied the 24 statements in a self-paced manner. Participants did not receive any information about whether a statement was true or false, although they were instructed that "some of them are true and some of them are false."

During the rating phase, participants rated the 24 target statements and the 12 foil statements in a random order. Half of the statements that were provided were purposefully awkwardly phrased to reduce fluency. The other half of the statements were clearer in phrasing. They were instructed to rate each statement at their own pace for truthfulness (true or false) and then how confident they were about their judgment (1-5). After the confidence rating, they saw the propositions for the statement. Each proposition

that composed the target statement was presented on a separate screen. They rated how plausible each proposition was on a scale of 1-4, as in Pilot study (See Table 1 for examples of the rating anchors). All judgment tasks were self-paced, and the response times while making truth, confidence, and plausibility judgments were recorded.

Results and Discussion

The main question was whether fluency or plausibility (or both) are likely to underlie the illusion of truth effect. I addressed this question by analyzing the plausibility ratings of the propositions and the time it took to make the judgments. I conducted two main sets of analyses: ANOVAs and a series of correlation and regression analyses.

ANOVAs. I conducted 2 (repetition) x 2 (clarity) x 2 (number of propositions) mixed ANOVAs on three main Dependent Variables: truth-confidence judgments (which range from -5 to +5) and on both reaction times (how long it took to make the truth judgment, and confidence judgment).

Veracity-Confidence Judgments. We expected veracity-confidence ratings to be influenced by clarity, repetition, and number of propositions. See Figure 8 for mean veracity-confidence ratings as a function of these variables. There was a main effect of repetition, $F(1,76) = 21.278, p < .001, \eta_p^2 = 0.218, MSE = 5.159$, where the repeated statements ($M = 1.277, SD = 1.325$) had higher ratings compared to new statements ($M = 0.433, SD = 1.207$). This is a replication of the Illusion of Truth Effect. As expected, there was also a main effect of clarity, $F(1,76) = 13.330, p < .001, \eta_p^2 = 0.149, MSE = 1.714$, such that the clear statements ($M = 4.191, SD = 3.982$) had higher ratings

compared to awkward statements ($M = 2.650$, $SD = 3.901$). Number of propositions, however, did not reach significance, $F(1,76) = 3.501$, $p = .065$, $\eta_p^2 = 0.044$, $MSE = 2.034$.

The main effects were qualified by several interaction effects. There were significant two-way interactions between repetition and clarity, $F(1,76) = 25.162$, $p < .001$, $\eta_p^2 = 0.249$, $MSE = 1.652$, repetition and proposition number, $F(1,76) = 5.695$, $p = .019$, $\eta_p^2 = 0.070$, $MSE = 2.297$, and clarity and proposition number, $F(1,76) = 25.051$, $p < .001$, $\eta_p^2 = 0.248$, $MSE = 1.632$. However, because there was also a significant three-way interaction, $F(1,76) = 12.330$, $p = .001$, $\eta_p^2 = 0.140$, $MSE = 2.201$, in lieu of following up on all statistically significant two-way interactions I followed up the three-way. See Table 3 for detailed follow up t -test results and Figure 8 for the means.

Statements that had three propositions showed repetition effects in that both clear and awkward statements were showing an inflated veracity-confidence judgment as would be seen in illusion of truth. For statements that had two propositions, only the awkward statements showed a repetition effect where the new had lower veracity confidence ratings compared to repeated statements. When there was less plausibility information to base the judgment on (because there were fewer propositions), fluency related information became a more reliable basis for judgment.

Reaction Time. The reaction time data were separated into two categories, as in Experiment 1: Veracity Rating Time (VRT) and Confidence Rating Time (CRT). Both reaction time data were positively skewed (see Figures 9 and 10) so I log-transformed the data for further analyses.

The fluency account predicts that veracity judgments should be sensitive to repetition, as processing is more fluent with more repetitions. For veracity judgment reaction time (see Figure 11 and Table 3), there was indeed a main effect of repetition, $F(1,76) = 48.997, p < .001, \eta_p^2 = .392, MSE = 0.018$, where the once repeated statements ($M = 5.798$ s, $SD = 1.457$ s) were significantly more quickly rated than new statements ($M = 6.896$ s, $SD = 1.529$ s). There was also a main effect of clarity, $F(1, 76) = 7.401, p = .008, \eta_p^2 = .089, MSE = 0.017$, where the clear statements ($M = 6.113$ s, $SD = 1.517$ s) were more quickly rated than awkward statements ($M = 6.534$ s, $SD = 1.468$ s), consistent with the fluency account. However, there was no main effect of proposition number, $F(1, 76) = 0.900, p = .346, \eta_p^2 = .012, MSE = .013$. Thus, unlike in Experiment 1, repetition and clarity both made veracity judgment times faster, which is what the fluency account predicts.

These results were further qualified by a number of interaction effects. There was a two-way interaction between repetition and clarity $F(1,76) = 5.621, p = .020, \eta_p^2 = 0.069, MSE = 0.016$, but not for repetition and number of propositions, $F(1,76) = 0.101, p = .751, \eta_p^2 = .001, MSE = .012$, or proposition number and clarity, $F(1,76) = 2.189, p = .143, \eta_p^2 = .028, MSE = .014$. Because the three-way interaction was also significant, $F(1,76) = 7.110, p = .009, \eta_p^2 = 0.086, MSE = 0.014$, I reported follow-ups to the three-way interaction in lieu of following up on all statistically significant two-way interactions. I computed t -tests comparing repeated to non-repeated for each level of awkwardness and proposition number in order to test for illusion of truth effects (see Table 3 for detailed follow up t -test results). For statements with three propositions, both

awkward and clear statements showed significant repetition effects in that the repeated statements had significantly higher veracity confidence judgments compared to the new statements. However, for statements that had two propositions, only the awkward statements showed a significant repetition effect. Repeated statements received high veracity-confidence ratings across all sentence types, regardless of length or awkwardness. A post-hoc one-way ANOVA⁸ on repeated items' veracity confidence judgment and veracity judgment time separately showed that there is no effect of proposition or clarity. In contrast, new statements were sensitive to fluency, with both awkwardness and length having some impact on veracity-confidence judgments.

For confidence reaction time (see Figure 12), there was no main effect of repetition, $F(1, 76) = 0.719, p = .399, \eta_p^2 = 0.009, MSE = 0.016$, no main effect of proposition number, $F(1, 76) = 3.288, p = .074, \eta_p^2 = 0, MSE = 0.012$ and no main effect clarity, $F(1, 76) = 0.718, p = .399, \eta_p^2 = 0.009, MSE = 0.014$. There were no two-way or three-way interactions (all $F_s < 1$).

Correlation and Regression Analyses. My next set of analyses all examined the predictive value of measures relating to fluency (mostly reaction time measures) and measures relating to plausibility (derived from the plausibility ratings of the propositions that made up the statements).

⁸ Dependent Variable: veracity confidence judgments; Independent Variable: proposition $F(1,1384) = 0.018, p = .894$; clarity $F(1,1384) = 0.851, p = .356$.
Dependent Variable: veracity judgment reaction time; Independent Variable: proposition $F(1,688) = 3.614, p = .058$; clarity $F(1,688) = 0.064, p = .801$.

Correlating Reaction Time and Plausibility Ratings with Veracity-Confidence

Judgments. My next question was whether fluency (reaction time) or plausibility (assessed by the plausibility ratings) was a better predictor of veracity-confidence judgments for individual statements. Therefore, I conducted a multiple regression analysis to evaluate whether reaction time or plausibility rating is a better predictor of veracity-confidence ratings. The correlation coefficients of the predictors (see Table 4) showed that all measures of plausibility ratings and reaction time measures while completing the plausibility rating were significantly correlated except for the veracity judgment reaction time ($r = -.047, p > .05$) and confidence judgment reaction time ($r = -.034, p > .05$).

As suggested by a committee member, I conducted a regression model that predicted veracity-confidence judgment using the veracity judgment reaction time and minimum plausibility rating. These measures were identified as the most theoretically-relevant to comparing fluency and plausibility. The fluency account would suggest that the time it took to make the veracity judgment is the best predictor whereas the plausibility account would suggest that the lowest plausibility rating of a proposition would be the best predictor. The results favored the plausibility account in that 17.2% of the variance was explained by the minimum plausibility rating, while confidence judgment reaction time explained no additional variance. This is unsurprising considering the non-significant correlations between veracity judgment reaction time and veracity-confidence judgments.

Best-Predictor Regression Analyses. To make a less conservative test of both accounts, I conducted two separate forward stepwise linear multiple regression models to find the best predictor from the reaction times (i.e., fluency) and the best predictor from the plausibility ratings. A third model compared the best reaction time variable and the best plausibility variable.

The first model was conducted to find the best reaction time⁹ predictor for truth judgments. See Table 5 for a summary of the regression coefficients. The dependent variable was the confidence-veracity judgment ratings and the predictors that were added in the model are as followed: Total RT (entire time it took to rate the plausibility ratings), Minimum RT (minimum time spent rating a particular proposition's plausibility ratings), Maximum RT (maximum time spent rating a particular proposition's plausibility ratings), Average RT (average time spent rating a particular proposition's plausibility ratings), VRT (time spent making the truth judgment) and CRT (time spent making the confidence judgment about their previous truth judgment).

The fluency account suggested that the best predictor will be reaction time during the veracity-confidence ratings (VRT), as that would form the basis for participants' sense of fluency for the statement's content (while piecing together the reaction time for proposition processing). However, for the plausibility account, the minimum reaction time while rating the plausibility for the propositions would be a better predictor. The results favored the plausibility account in that the best and only predictor was minimum

⁹ All reaction time data was log-transformed for the analysis. Doing a quadratic analysis of the reaction time did not change the results of models 1 and 3.

reaction time for rating a particular proposition's plausibility rating ($\beta = -.086, p < .001$), though it explained only 0.6% of the variance.

The second model was conducted to find the best plausibility ratings predictor for truth judgments. See Table 6 for a summary of the regression coefficients. The dependent variable was the confidence-veracity judgment ratings and the predictors that were added in the model are as followed: Average Plausibility (average plausibility ratings for the propositions associated with the target statement), Minimum Plausibility (minimum plausibility rating for a particular proposition associated with the target statement), and Maximum Plausibility (maximum plausibility rating for a particular proposition associated with the target statement). The plausibility account would predict the best predictor to be the minimum plausibility value for the proposition rating. The fluency account wouldn't predict that the rating value for propositions would matter as much but rather the length or number of propositions, which is why fluency would predict a reaction time measure would be a better predictor. The results showed that the minimum plausibility rating explained 7.6% of the variance, and was the best predictor for truth judgments, again favoring the plausibility account.

The third model was conducted to understand the relationship between fluency and plausibility by entering the best predictors for reaction time (i.e., Minimum Reaction Time) and plausibility ratings (i.e., minimum plausibility rating) determined in the previous models. See Table 7 for a summary of the regression coefficients. I predicted that the total reaction time will be more predictive of the illusion of truth from the fluency perspective. But the minimum plausibility rating would be a better predictor than any of

the reaction time measures. The results were in favor of the plausibility account. Once minimum plausibility was accounted for, reaction time had no additional predictive power.

Effects of Statement Clarity. As part of the post-hoc analyses, to better understand the data, I looked at the regression coefficients separately by statement clarity. For awkward statements, the stepwise regression analysis showed that minimum plausibility ratings explained 11.4% of the variance, $R^2 = .114$, $F(1, 1384) = 178.946$, $p < .001$. For clear statements, the stepwise regression analysis showed that minimum plausibility rating explained 4.1% of the variance, $R^2 = .041$, $F(1, 1383) = 58.663$, $p < .001$ and adding the minimum reaction time increased 0.6%, $\Delta R^2 = .006$, $F(1, 1383) = 9.327$, $p = .002$. When running a stepwise regression on Veracity-Confidence Judgments for just the statements that were judged true, the minimum plausibility explained 3.9% of the variance, $R^2 = .039$, $F(1, 1858) = 75.228$, $p < .001$, while the reaction time was not significant.

Conclusion

In this experiment, I tested predictions from the plausibility and fluency account and found that plausibility of statements' propositions were stronger predictors of the truth judgments than reaction time measures. This does not mean fluency did not matter. Rather, when strong plausibility exists, it may be more valued than fluency. The regression analyses on the variables suggest the time it takes to judge the lowest plausibility rating is the best predictor within the reaction time domain, but the actual value of the lowest plausibility rating is a better predictor. The ANOVA analysis seems

to confirm what was already known about the Illusion of Truth effect in that repetition does influence subsequent veracity-confidence judgments. The post-hoc analyses showed that repetition also increases mean plausibility ratings of the propositions. The regression analysis shed more light into the role of plausibility and rating time on veracity-confidence judgments. And as the plausibility account would predict, plausibility ratings (either minimum or average) was a better predictor of veracity-confidence judgments compared to the time it took to rate veracity.

The fluency account received support as well, with awkward statements rated less true than clear statements. The plausibility account does not predict much of a difference between the two phrasings, as plausibility would mainly be affected by the content. The 3-way interaction can be interpreted as little effect of fluency factors (i.e., clarity, length, or repetition) with repetition, but a reliable effect without. It is hard to tell if more repetitions and delayed judgment would make fluency stronger.

For reaction time, the fluency and plausibility accounts each predict a main effect of repetition, clarity, and number of propositions. The fluency account further predicts a main effect of clarity, because the awkward and wordy statement is processed less easily and therefore will likely take longer to process than clearer statements. By the same logic, fluency predicts a main effect of number of propositions. However, the reaction time data show a three-way interaction of repetition, clarity, and proposition number, a main effect of repetition only and a two-way interaction between clarity and repetition. The new statements took longer to make veracity judgments overall, specifically when the statements were awkwardly worded, which supports the fluency account.

CHAPTER III

GENERAL DISCUSSION

The goal of the present experiments was to test predictions derived from the plausibility account of the illusion of truth effect. The fluency account proposes that repetition causes ease of processing, which then inflates veracity judgments. In contrast, the plausibility account suggests that highly available content is strengthened through repetition, and that people rapidly compare statements with what is readily available in memory, inflating veracity judgments with repetition. Experiment 1 showed that there is an interaction between content valence and repetition frequency, consistent with the plausibility account. Experiment 2 showed that plausibility ratings of propositions were a better predictor of veracity confidence judgments than reaction time. Taken together, the results suggested that the plausibility account provided explanation for the data, and although fluency still contributed to truth judgments, ease of processing may not be the most important factor in producing the illusion of truth effect in all cases.

Assumptions of the Plausibility Account

I have not spelled out in detail how plausibility functions, but any model based on plausibility needs to do two things: it has to match the target statement to information in memory (the match phase), and it needs to rapidly compare the information in memory to the statement to judge plausibility (the check phase). Similar to Reder's (1982) plausibility account of sentence verification paradigm, I proposed that memory models

can shed some light on likely assumptions of a more detailed plausibility theory for illusion of truth effect (e.g., ACT-R: Anderson, 2005; SAM: Atkinson & Shiffrin, 1968; Gillund & Shiffrin, 1984). In particular, I speculated that it may work much like familiarity-based recognition memory, except instead of evidence accumulating that the item was seen before, evidence accumulates about the plausibility of the item.

Plausibility accounts imply that a full match between the statement and relevant information in memory is not necessary to obtain the illusion of truth effect. For the match stage, the information is triggered quickly and in greater detail if the targets are more (a) temporally recent, (b) strongly associated, and (c) overlapping in content with the cue. These assumptions are all true of recognition memory based on familiarity as well. I assume accumulated evidence of plausibility, which is why it predicts an increase in plausibility after repetition. During the check stage, people quickly detect contradictions or confirmations based on what they know about the target, once a match is made (i.e., after the match stage). If the evidence confirms or contradicts the statement, people will make the judgment as *true* or *very unlikely to be false/very likely to be true*. If there is insufficient evidence or disconfirming evidence, people will make the judgment as *False* or *very unlikely to be true/very likely to be false*.

The choice to respond either true or false in the check stage depends on whether the information in memory contradicts or confirms the statement, and whether it is considered sufficient. Like recognition memory, evidence accumulates, but in recognition judgments there is only evidence “for” the item being presented before, whereas in plausibility, the evidence is whether the evidence is “consistent” with the item being

presented or “inconsistent.” Thus, evidence accumulates both for statements being plausible and statements being implausible, not just evidence for the item being old. Consistent with the idea that we accumulate evidence of inconsistency, when asked to make truth judgments for statements whose content has changed from the initial exposure, participants were quick to correctly give a false judgment instead of a faulty and inflated truth judgment (Unkelbach & Rom, 2017). Furthermore, it is probably crucial that the participants notice that there was a change in the statement, as found in recursive reminding literature. When participants are given word pairs A-B at initial study and are presented with A-D at test, those who detected a change in the word pair showed accurate recognition (Wahlheim & Jacoby, 2013).

Post Hoc Analysis Results

Though the results of the proposed analyses generally provided evidence in favor of the plausibility account’s predictions, I conducted some post-hoc analyses to answer some additional questions about plausibility and fluency accounts.

How comparable is veracity confidence to recognition confidence? The plausibility account assumes a familiarity-based recognition type judgment. When confidence is lower, the time it takes to make recognition judgments is usually longer (for reviews, see Hockley, 2008; Kahana & Loftus, 1999). To test this assumption, I conducted a bivariate correlation analysis with the confidence judgment and confidence judgment reaction time. The confidence judgment was calculated without the veracity values which just reflected the intensity of the confidence. The confidence judgment reaction time was the log transformed data that was used in previous analyses. The results

found no significant correlation¹⁰ between confidence judgments and reaction time, $r(942) = -.019, p = .566$. This is in line with the correlation results found with the veracity-confidence judgments in Experiment 2, where both the veracity judgment reaction time and confidence judgment reaction time were statistically nonsignificant. Confidence in veracity judgments does not appear to work exactly the same way as familiarity-based recognition judgments do.

What does it mean to “easily process” information in fluency? A skeptical reader might argue that fluency is still the best explanation for our results, and that plausibility is nothing but renaming it. Fluency is a construct that depicts the ease with which information is processed (e.g., Alter & Oppenheimer, 2009). However, the definition of fluency is unclear in the literature: what is ease? Without a clear operational definition of processing ease, Unkelbach and Rom (2017) came closest to finding a concrete definition of it: response time. The more fluently an item is processed, the quicker the response.

However, contrary to this interpretation, they found that participants were able to make very quick “false” responses. If fluency were the only driving factor in truth judgments, then this result should not have been possible (c.f., Unkelbach & Rom, 2017). Similarly, when pre-exposed sentences were contradictory to the target statement despite the fluent lexical processing, participants showed an illusion of false effect in that their truth judgments decreased more so that they thought it to be false. In Experiment 1 of the current study, repeated information that contradicted the target statement inflated the

¹⁰ An ROC curve analysis showed the same non-significant results. And the results were the same for Experiment 1 $r(84) = -.091, p = .408$.

false judgment, inducing an “illusion of falsehood” effect. This would not have been possible if the sentence content was simply available and fluent. Participants would have to compare what was exposed to them and what the target statement said.

A fluency theorist could argue that reaction time reflects only “perceptual fluency” whereas “conceptual fluency” reflects a general sense of ease, not just reaction time. In typical conceptual fluency studies in recognition, ideas are primed using related items, and therefore processed more easily (e.g., Jacoby & Dallas, 1981; Lanska, Olds, & Westerman, 2014). Priming increases both hits and false alarms in recognition tests. Further, the notion of conceptual fluency would have to be expanded to include a kind of anti-conceptual fluency, whereby items also sometimes seem more *false* when they are highly familiar. It is difficult to see how a priming mechanism alone could account for this pattern, which suggests content information is necessary for making veracity judgments. In Experiment 1 of the current study, the contradicting pre-exposed sentences were repeated and had similar lexical familiarity to the statements, but the veracity confidence judgments decreased with more repetitions. Participants were able to determine the veracity by examining the content congruency between pre-exposed sentences and the target statement.

How is the referential theory different from plausibility? The relationship between plausibility and the referential theory (Unkelbach & Rom, 2017; Unkelbach et al., 2019) should be further examined. For instance, are conceptually fluent items higher in reference connection and therefore are judged to be plausible?

Unkelbach and colleagues (2019) discussed a referential theory that suggests the number of references between items inflates the truth judgment. When participants were asked to think about the relationship between the statements, the truth judgments increased as opposed to when they were asked to simply read the statements or report which side of the screen the statements were presented on. This finding was discussed in terms of greater connection between the references, which the plausibility account would suggest is the increased strength and detection in the match phase.

Could the plausibility measure be reactive? Within the current studies, the veracity judgment time differed for Experiment 1 (near 2 s) and Experiment 2 (near 6 s). Does adding an instruction of subsequent plausibility rating increase the likelihood of spending more time to think through the statement? Previous literature has found that rating pre-experimental familiarity (e.g., *What/How much do I know about Ernest Hemingway?*) to the target statement (e.g., *Ernest Hemingway was an ambulance driver during World War I*) affected the truth judgment (Schwartz, 1984). Simply seeing the topic (e.g., *Statue of Liberty*) related to the target statement (e.g., *The Statue of Liberty faces Southeast*) also affected the truth judgment (e.g., Begg et al., 1985). This is consistent with the possibility that giving the instructions to participants that they will be asked to judge the plausibility of the parts of the statements could affect the veracity judgment time. Future studies should separate assessment of plausibility and veracity-confidence, perhaps by collecting veracity-confidence judgments first without warning people that plausibility will be assessed separately after a delay.

Implications and Future Directions

The current study investigated the value of plausibility in explaining the illusion of truth effect. However, I believe plausibility could potentially partially explain many more phenomena. With “fake news” in the media – especially on social media – receiving so much attention lately, understanding the role of information consumers’ plausibility judgments in consuming the news has never seemed more important.

One reason fake news may be shared so often is that it is designed in such a way as to meet peoples’ plausibility threshold, by fitting well with their pre-existing beliefs. If so, this might lead consumers to spread the news to other people and thereby increase the plausibility of the false information to yet more people. In Experiment 1, we see the baseline veracity confidence judgments on new statements (0-repetitions in Figure 2) is different between the content valence conditions. In recognition memory, the threshold for deciding items were old changes depending on the strength of the evidence (e.g., Hirshman, 1995), with stronger available evidence leading people to adopt stricter criteria for calling items old. Analogously, in illusion of truth, we might expect peoples’ threshold to depend on the relative availability of information, such that we set a stricter criterion for calling something true when the available evidence tends to be strong. Consistent with this view, in Experiment 1, belief in the new statements was lowest in the exact repetition condition. This is the condition where the best match occurs, leading to the strongest evidence for the old statements being true. However, when pre-exposed sentences that had content valence, even when participants saw the statement for the first time, the veracity confidence judgments were high relative to the exact condition.

Without proper time and effort to scrutinize the information, the baseline level of plausibility presented in a news story could influence the subjective truthfulness of the rest of the story to a person, encouraging the spread of false information. Especially on social media, such decisions may be made quickly and rely on plausibility and fluency rather than sustained thought. It is worth further investigating the role of System 1 versus System 2 processing (e.g., Kahnemann, 2011) in social media sharing decisions.

A similar point is that one's background and prior beliefs are likely to influence the balance of fluency and plausibility in judging a statement. It is no secret that people tend to favor information that confirms their existing beliefs (e.g., Wason, 1968). In some cases, personality traits might even affect our plausibility judgment mechanisms. For example, people with high levels of schizotypy were more likely to judge a delusional statement to be true compared to low-schizotypy controls (e.g, Moritz et al., 2012) – because delusional content makes sense to them and seems more plausible.

Another aspect that can be elaborated on is the relationship between plausibility and initial attitudes towards a statement. In the sleeper effect (e.g., for meta-analysis see Kumkale & Albarracín, 2004), where decrease in source memory after a delay increases the persuasiveness of the initially discounted information. For example, a statement like “iron is atomic number 26” was presented with an uncredulous source (e.g., a student who started learning Latin) which initially would make the statement less persuasive. However, at a delay, participants forgot who told the statement, but remembers the statement itself which makes them feel like it is persuasive. Plausibility account naturally predicts this phenomenon. Plausibility account assumes that temporally recent target,

additional information, and cue associations are more likely to be matched easily. The statement and source is highly associated with the cue immediately; however, as recency fades, statement ceases to strongly associate its source. Therefore, plausibility will predict how the participants were likely to judge a statement's lack of persuasiveness immediately after the presentation of statement and uncredulous feedback. After the delay the statement is judged to be more persuasive because the cue (i.e., statement or parts of the statement) and statement share overlapping content, which therefore remains a good match, even though the source information is no longer readily available.

Plausibility could also be a key mechanism we use to decide whether a statement requires elaboration or not. For example, a statement's agreeableness is rated based on the strength of the argument that supports it (e.g., see Elaboration Likelihood Model by Petty & Cacioppo, 1980). When the argument strength is left ambiguous and there is no other manipulation of strength (e.g., repetition), may affect the participants' ability or motivation to check the plausibility of the arguments will determine their judgments about the statements. Recently, a study showed that individual differences (i.e., cognitive ability and need for cognition) influenced the illusion of truth effect (De Keersmaecker et al., 2019). Need for cognition might help determine the amount of processing time devoted to plausibility checking. Similarly, one study found that lack of cognitive ability, induced by a cognitive load manipulation increased the likelihood of making illusory truth judgments (e.g., Skurnick, Yoon, Park, & Shwartz, 2005). The more highly motivated participants were to think about the statements before making the judgments, the less susceptible they were to the illusion of truth effect (Garcia-Marques & Silva,

2017). There was an overall illusion of truth effect in that repeated statements were judged to be truer compared to new statements. However, the effect size decreased when the participants in the no-load condition were instructed to be as accurate as possible. Without manipulation of the cognitive load, studies looking at people's predispositions to emphasize their feelings in most contexts (Sunbar, Kardes, & Wright, 2015) or to be skeptical (Difonzo et al., 2016) found similar results. Participants who utilized their feelings more (i.e., high on the Need for Affect – a measure of one's need to focus on their feelings in most contexts as opposed to using logic and careful processing) when making truth judgments were more influenced by the repetition effect, in that repeated statements were judged to be more truthful. However, for participants who were low on the Need for Affect scale, there were no illusion of truth effect found (Sunbar et al., 2015). Participants low on dispositional skepticism were more affected by more frequent repetition whereas participants high in skepticism were less affected (Difonzo et al., 2016). In sum, it seems the illusion of truth effect could decrease with careful and thoughtful processing of the statement before judgment. However, when participants are deprived of this cautious thinking, there was room for error in the truth judgment because having a sense of plausibility is enough for them to inflate the veracity of the statement.

In conclusion, this paper proposed that a plausibility mechanism should be added onto the current fluency account of illusion of truth. The effect size for perceptual fluency is rather small, and there is still variance left that needs further explanation beyond perceptual fluency. Robust effects such as the illusion of truth effect are often influenced

by multiple mechanisms rather than having one driving factor. The current paper suggests that plausibility could be one of these mechanisms in illusion of truth.

REFERENCES

- Anderson, J.R. (1974). Retrieval of propositional information from longterm memory. *Cognitive Psychology*, *5*, 451–474.
- Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036-1060.
- Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*, 89-195.
- Arkes, H.R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, *2*, 81-94.
- Arkes, H.R., Boehm, L. E., & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology*, *26*, 576-605.
- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*, 219-235.
- Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. Adequacy of the repetition account. *Journal of Experimental Psychology: Human, Learning and Memory*, *5*, 241-252.
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source, recollection, statement familiarity and illusion of truth. *Journal of Experimental Psychology: General*, *121*, 446-458.
- Brennan, S. E., and M. Williams. 1995. The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*, 383-398.
- Brown, A. S., & Marsh, E. J. (2009). Creating illusions of past encounter through brief exposure. *Psychological Science*, *20*, 534-538.
- Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1088–1100.

Bruett, H. & Leynes, P. A. (2015). Event-related potentials indicate that fluency can be interpreted as familiarity. *Neuropsychologia*, 78, 41-50
Boehm, L. E. (1994). The Validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, 20, 285-293.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, 106, 265-289.

Dechêne, A., Stahl, S., Hansen, J., & Wänke, M. (2009). Mix me a list: Context moderates the truth effect and the mere-exposure effect. *Journal of Experimental Social Psychology*, 45, 1117-1122.

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

Difonzo, N., Beckstead, J. W., Stupak, N., & Walders, K. (2016). Validity judgments of rumors heard multiple times: The shape of the truth effect. *Social Influence*, 11, 22-39.

Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, 20, 540-551.

Faul, F., Erdfelder, E., Lang, A. G., Buchner, A. (2007). G*Pwer 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

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*, 1, 993-1002.

Forster, M., Fabi, W., & Leder, H. (2015). Do I really feel it? The contributions of subjective fluency and compatibility in low-level effects on aesthetic appreciation. *Frontiers in Human Neuroscience*, 9, 1-11.

Garcia-Marques, T., Silva, R. R., Reber, R., & Unkelbach, C. (2015). Hearing a statement now and believing the opposite later. *Journal of Experimental Social Psychology*, 56, 126-129.

Garcia-Marques, T., Prada, M., & Mackie, D. M. (2016). Familiarity increases subjective positive affect even in non-affective and non-evaluative contexts. *Motivation and Emotion*, 40, 638-645.

Garcia-Marques, T., Silva, R. R., & Mello, J. (2016). Judging the truth-value of a statement in and out of a deep processing context. *Social Cognition*, 34, 40-54.

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology, 44*, 687-691.

Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: The impact of linguistic concreteness and level of construal on subject truth. *Personality and Social Psychology Bulletin, 36*, 1576-1588.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*, 107-112.

Henkel, L. A., & Mattson, M. E. (2011). Reading is believing: The truth effect and source credibility. *Consciousness and Cognition, 20*, 1705-1721.

Hirshman, E. (1995). Decision processes in recognition memory: criterion shifts and the list-strength paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 292-313.

Hockley, W. E. (2008). Memory search: A matter of time. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference – Vol. 2* (pp. 417-444). Amsterdam, The Netherlands: Academic Press.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Experimental Psychology, 56*, 326-338.

Johar, G. V., & Roggeveen, A. L. (2007). Changing false beliefs from repeated advertising: The role of claim-refutation alignment. *Journal of Consumer Psychology, 17*, 118-127.

Jiang, Y., & Hong, J. (2014). It feels fluent, but not right: The interactive effect of expected and experienced processing fluency on evaluative judgment. *Journal of Experimental Social Psychology, 54*, 147-152.

Kahana, M. J., & Loftus, G. (1999). Response time versus accuracy in human memory. In R. J. Sternberg (Ed.), *The nature of cognition* (pp. 322-384). Cambridge, MA: MIT Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Koch, A. S., & Forgas, J. P. (2012). Feeling good and feeling truth: The interactive effects of mood and processing fluency on truth judgments. *Journal of Experimental Social Psychology, 48*, 481-485.

- Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin*, 130, 143-172.
- Killer, J. K., Lloyd, M. E., & Westerman, D. L. (2008). When does modality matter? Perceptual versus conceptual fluency-based illusions in recognition memory. *Journal of Memory and Language*, 58, 1080-1094.
- Kurilla, B. P. & Westerman, D. L. (2008). Processing fluency affects subjective claims of recollection. *Memory & Cognition*, 36, 82-92.
- Koch, T., & Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication*, 63, 933-1010.
- Kelley, C.M. & Lindsay, D.S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, 32, 1-24.
- Kelley, C., M. & Rhodes, M. G. (2002). Making sense and nonsense of experience: Attributions in memory and judgment. *The Psychology of Learning and Motivation*, 41, 293-320.
- Lanska, M., Olds, J. M., & Westerman, D. L. (2014). Fluency effects in recognition memory: Are perceptual fluency and conceptual fluency interchangeable? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1-11.
- McGlone, M. S., & Tofiqbakhsh, J. (2000). Birds of a feather flock conjointly: Rhyme as reason in aphorisms. *Psychological Science*, 11, 424-428.
- Moreland, R. L., & Topolinski, S. (2010). The mere exposure phenomenon: A lingering melody by Robert Zajonc. *Emotion Review*, 2, 329-339.
- Moons, W. G., Mackie, D. M., & Garcia-Marques, T. (2009). The impact of repetition-induced familiarity on agreement with weak and strong arguments. *Journal of Personality and Social Psychology*, 96, 32-44.
- Moritz, S., Kother, U., Woodward, T. S., Veckenstedt, R., Dechêne, A., & Stahl, S. (2012). Repetition is good? An internet trial on the illusory truth effect in schizophrenia and nonclinical participants. *Journal of Behavior Therapy and Experimental Psychiatry*, 43, 1058-1063.
- Nadarevic, L., & Erdfelder, E. (2014). Initial judgment task and delay of the final validity – rating task moderates the truth effect. *Consciousness and Cognition*, 23, 74-84.

- Ozubko, J. D., & Fugelsang, J. (2011). Remembering makes evidence compelling: Retrieval from memory can give rise to the illusion of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 270-276.
- Ozubko, J. D., & Joordens, S. (2007). The mixed truth about frequency effects on free recall: Effects of study list composition. *Psychological Bulletin & Review*, *14*, 871-876.
- Parks, C. M., & Toth, J. P. (2006). Fluency, familiarity, aging and the Illusion of truth. *Aging, Neuropsychology, and Cognition*, *13*, 225-253.
- Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: William C. Brown.
- Roggeveen, A. L. & Johar, G. V. (2002). Perceived source variability versus familiarity: Testing competing explanations for the truth effect. *Journal of Consumer Psychology*, *12*, 81-91.
- Reber, R. & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*, 338-342.
- Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, *89*, 250-280.
- Santos, A. S., Ramos, T., Garcia-Marques, L. & Carneiro, P. (2017). “To-be-forgotten” statements become less true: Memory processes involved in selection and forgetting lead to truthfulness changes of ambiguous sentences. *Scandinavian Journal of Psychology*, *58*, 205–210.
- Scholl, S. G., Griefeneder, R., & Bless, H. (2014). When fluency signals truth: Prior successful reliance on fluency moderates the impact of fluency on truth judgments. *Journal of Behavioral Decision Making*, *27*, 268-280.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195-202.
- Shapiro, S. (1999). When an ad’s influence is beyond our conscious control: Perceptual and conceptual fluency effects caused by incidental ad exposure. *Journal of Consumer Research*, *26*, 16-36.
- Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*, 713-726.

- Sunbar, A., Kardes, F. R., & Wright, S. A. (2015). The influence of repetitive health messages and sensitivity to fluency on the truth effect in advertising. *Journal of Advertising, 44*, 375-387.
- Topolinski, S., & Reber, R. (2010). Gaining insight into the “aha” experience. *Current Directions in Psychological Science, 19*, 402-405.
- Topolinski, S., & Strack, F. (2009). The analysis of intuition: Processing fluency and affect in judgements of semantic coherence. *Cognition and Emotion, 23*, 1465–1503.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232.
- 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*, 219-230.
- Unkelbach, C., Bayer, M., Alves, H., Koch, A., & Stahl, C. (2011). Fluency and positivity as possible causes of the truth effect. *Consciousness and Cognition, 20*, 594-602.
- Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology, 79*, 78-86.
- Unkelbach, C., Koch, A., Silva, R.R., & Garcia-Marques, T. (2019). Truth by repetition: Explanations and implications. *Current Directions in Psychological Science, 28*, 247-253.
- Unkelbach, C., & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition, 18*, 22-38.
- Unkelbach, C., & Rom, S. C. (2017). A referential theory of the repetition-induced truth effect. *Cognition, 160*, 110-126.
- Vosough, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*, 1146-1151.
- Wahlheim, C. N., & Jacoby, L. L. (2013). On importance of looking back: The role of recursive reminders in recency judgments and cued recall. *Memory & Cognition, 41*, 625-637.

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

Wänke, M., & Hansen, J. (2015). Relative processing fluency. *Current Directions in Psychological Science*, *24*, 195-199.

Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273-28.

Westerman, D., L., Lenska, M., & Olds, J. M. (2015). The effect of processing fluency on impressions of familiarity and liking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 426-438.

Whittlesea (2002). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General*, *131*, 96-115.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1-27.

Zürn, M., & Topolinski, S. (2017). When trust comes easy: Articulatory fluency increases transfers in the trust game. *Journal of Economic Psychology*, *61*, 74-86.

APPENDIX A
TABLES AND FIGURES

Table 1
Anchors and Questions used for Judgment Types in all Experiments.

Judgment Type	Question	Anchors
Truth	Is the statement true or false?	<ol style="list-style-type: none"> 1. True 2. False
Confidence	How confident are you about your previous response?	<ol style="list-style-type: none"> 1. I guessed. 2. I think maybe. 3. I think so 4. I'm pretty sure. 5. I know the answer.
Plausibility	How plausible do you think other UNCG students would rate the statement? (Pilot) How plausible do you think each proposition is? (E2)	<ol style="list-style-type: none"> 1. Very implausible 2. Implausible 3. Plausible 4. Very plausible.

Table 2

Experiment 1 t-tests for Pre-exposure Sentence Category Differences by Repetition for Confidence-Veracity Judgments (CVJ).

Pre-exposed content valence	0 vs. 1	0 vs. 2	1 vs. 2
Exact	-5.004 ^{***}	-4.900 ^{***}	-0.889
Confirming	-3.515 ^{**}	-3.502 ^{**}	0.402
Neutral	-0.448	0.369	0.034
Contradicting	0.896	2.140 [*]	2.040 ^a

Note. All comparisons' degrees of freedom were 20. ^{*} $p < .05$, ^{**} $p < .01$, ^{***} $p < .001$, ^a $p = .055$

Table 3

Experiment 2 t-tests of Confidence-Veracity Judgments, Veracity Judgment Reaction Time for Once Repeated – No Repetition by Clarity and Proposition Number.

Proposition	Clarity	CVJ	VJR
2	Clear	0.724	-1.002
	Awkward	7.314***	7.497***
3	Clear	1.708	4.070***
	Awkward	2.480*	4.022***

Note. All comparisons' degrees of freedom were 76. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4

Experiment 2 Correlation Coefficient Matrix.

	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. CVJ	.267* **	.276* **	.148* **	-.058* *	-.072* **	-.076* **	-.088* **	-.047	-.034
2. Ave. Plausibility	1	.851* **	.783* **	-.187* **	-.157* **	-.171* **	-.186* **	-.049	-.077 *
3. Min. Plausibility		1	.417* **	-.267* **	-.165* **	-.212* **	-.213* **	-.055	-.094 **
4. Max. Plausibility			1	-.009	-.080* **	-.037	-.066* *	-.035	-.002
5. Total RT				1	.553** *	.810** *	.778** *	.272* **	.239* **
6. Min. RT					1	.636** *	.827** *	.256* **	.274* **
7. Max. RT						1	.948** *	.284* **	.240* **
8. Ave. RT							1	.293* **	.271* **
9. VRT								1	.316* **
10. CRT									1

Note: * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5

Regression Coefficients for Reaction Time Data.

	Variable	B	SE B
Step 1	Constant	2.013	0.232
	Minimum Reaction Time	-1.307	0.265
	Adjusted R^2	0.006	
	F	6.798**	

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 6

Regression Coefficients for Plausibility Data.

	Variable	B	SE B
Step 1	Constant	-1.451	0.175
	Min. Plausibility	0.926	0.061
	Adjusted R^2	0.076	
	F	228.209***	
Step 2	Constant	-2.212	0.285
	Min. Plausibility	0.0591	0.116
	Ave. Plausibility	0.536	0.158
	Adjusted R^2	0.079	
	F	120.264***	
	ΔR^2	0.004	
Step 3	Constant	-2.037	0.294
	Min. Plausibility	0.27	0.179
	Ave. Plausibility	1.289	0.357
	Max. Plausibility	-0.471	0.2
	Adjusted R^2	0.081	
	F	82.156***	
	ΔR^2	0.002	
Step 4	Constant	-2.052	0.294
	Ave. Plausibility	1.787	0.134
	Max. Plausibility	-0.701	0.13
	Adjusted R^2	0.08	
	F	122.046***	
	ΔR^2	-0.001	

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$

Table 7

Regression Coefficients for Plausibility and Reaction Time Data.

	Variable	B	SE B
Step 1	Constant	-1.451	0.175
	Minimum Plausibility	0.926	0.061
	Adjusted R ²	0.076	
	F	228.209***	

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$

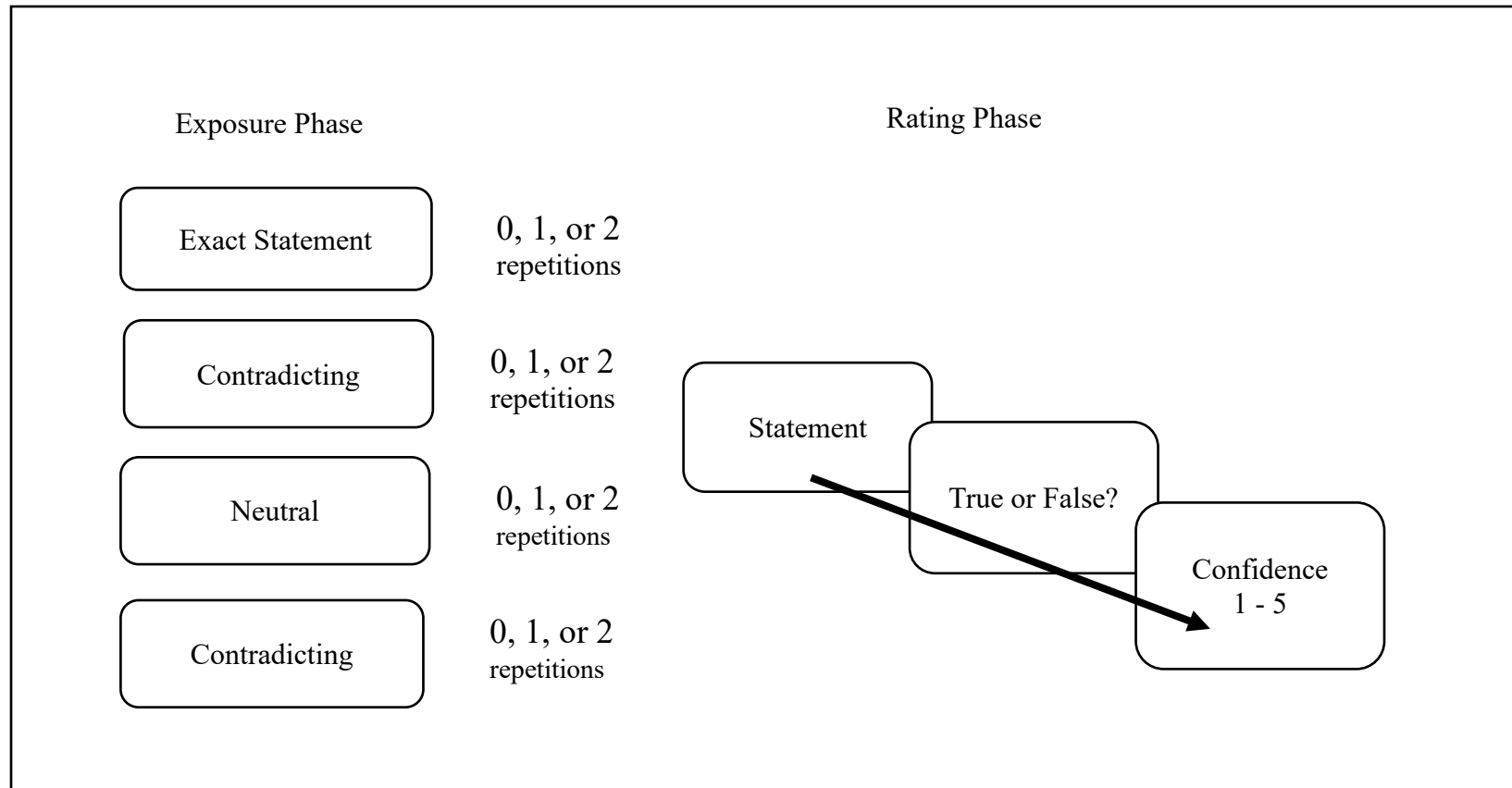


Figure 1. Diagram of Experiment Procedure.

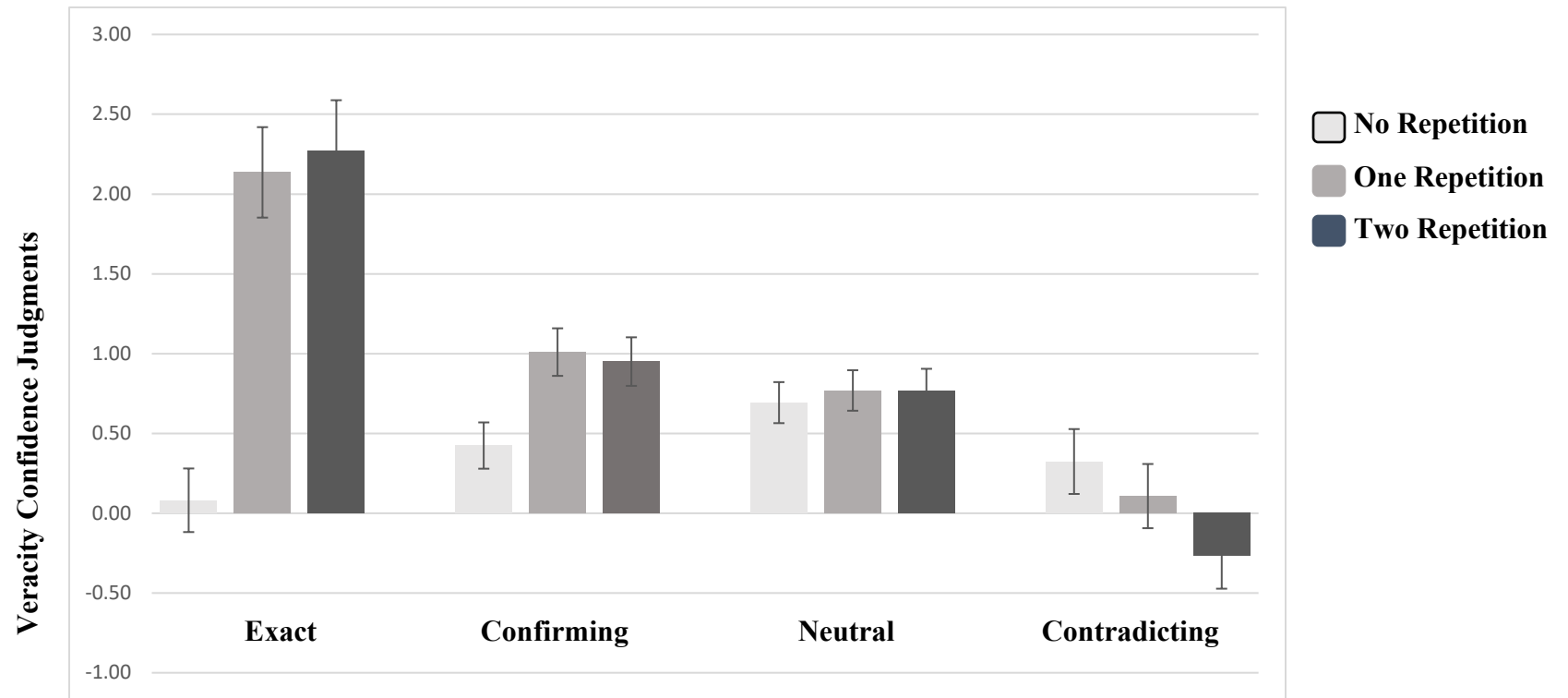


Figure 2. Experiment 1 ANOVA Results of Confidence-Veracity Judgments by Repetition for Pre-exposed Sentences.

Notes. The error bars indicate the standard error of the mean.

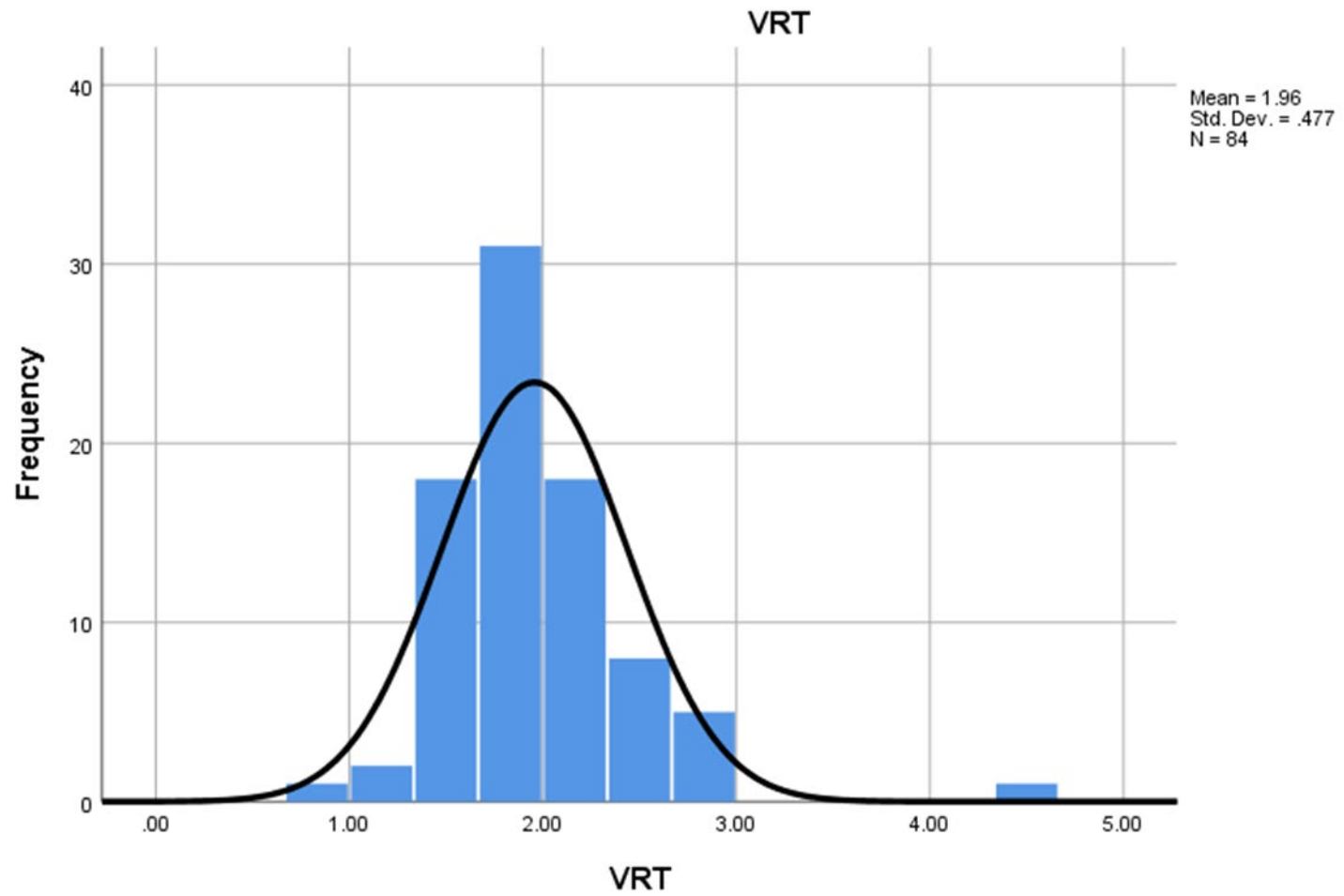


Figure 3. Histogram of Reaction Time Frequency for Experiment 1 Veracity Judgment Reaction Time.

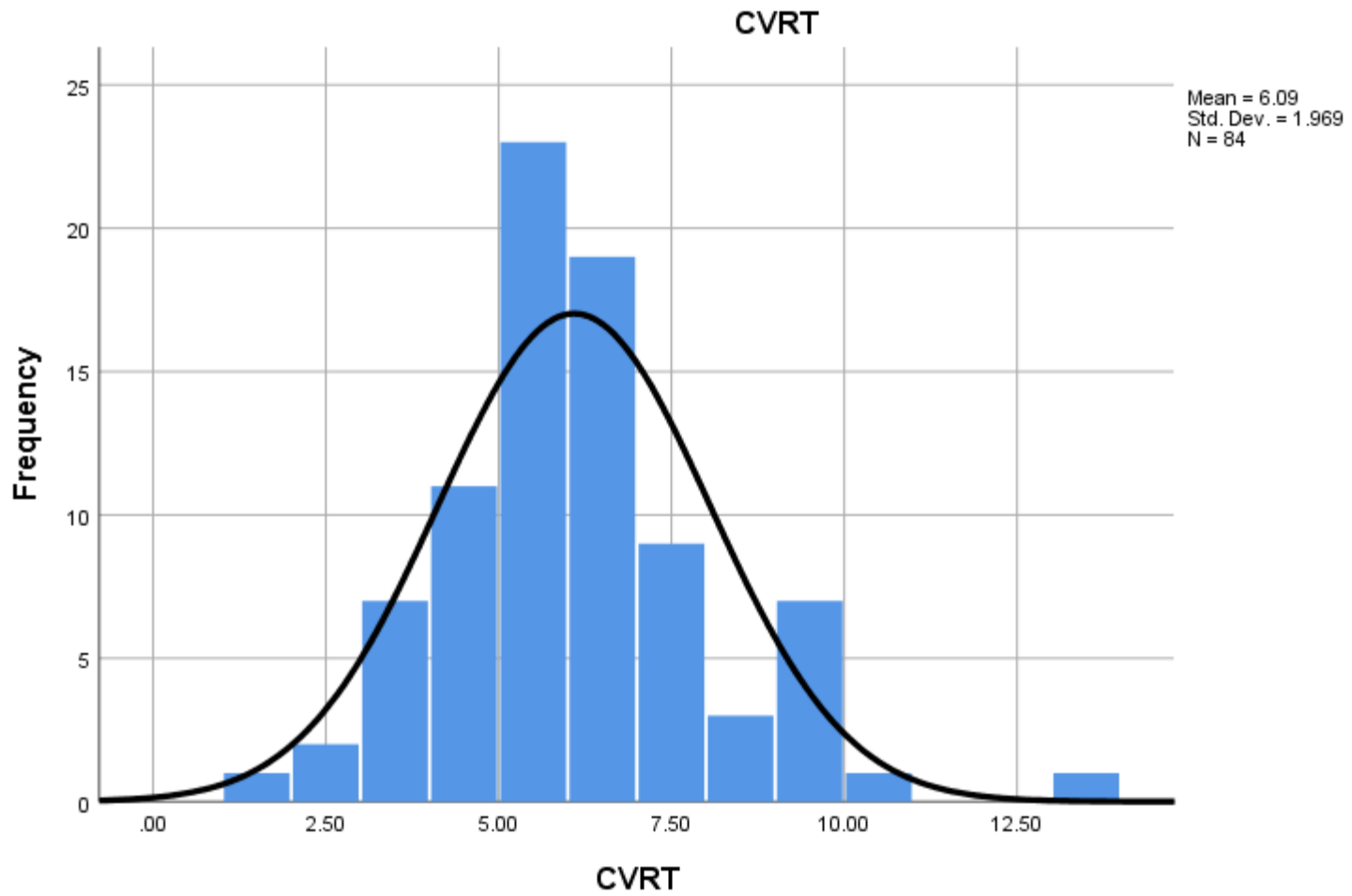


Figure 4. Histogram of Reaction Time Frequency for Experiment 1 Confidence Judgment Reaction Time.

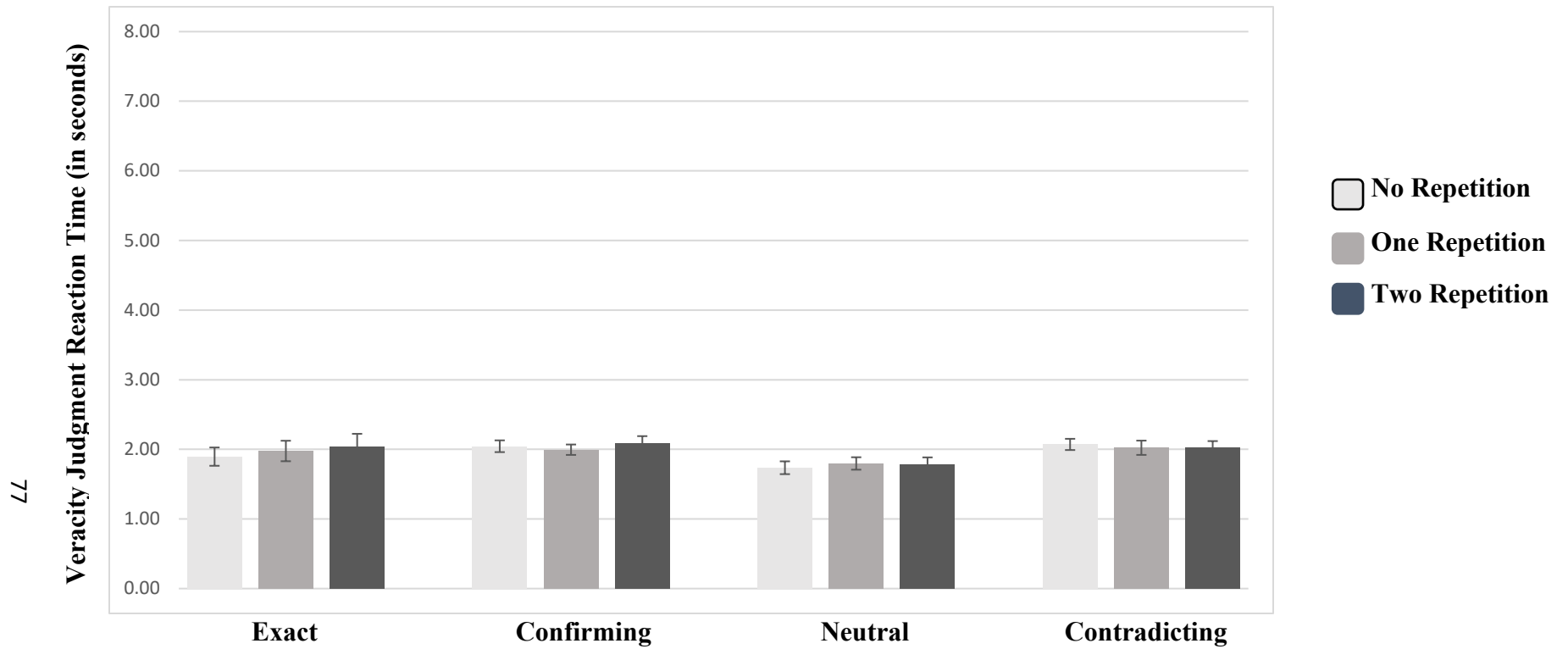


Figure 5. Experiment 1 ANOVA Results of Veracity Judgment Reaction Time by Repetition for Pre-exposed Sentences.

Notes. The error bars indicate the standard error of the mean.

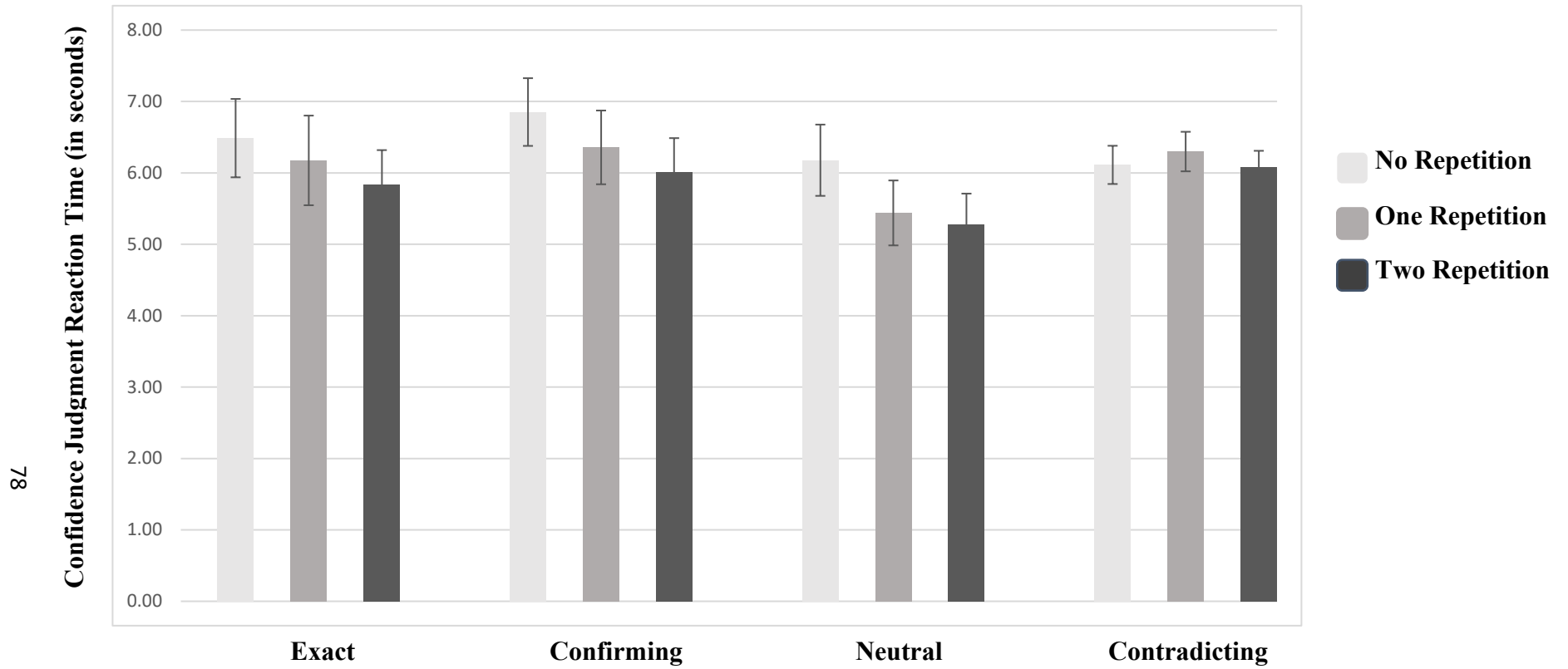


Figure 6. Experiment 1 ANOVA Results of Confidence Judgment Reaction Time by Repetition for Pre-exposed Sentences.

Notes. The error bars indicate the standard error of the mean.

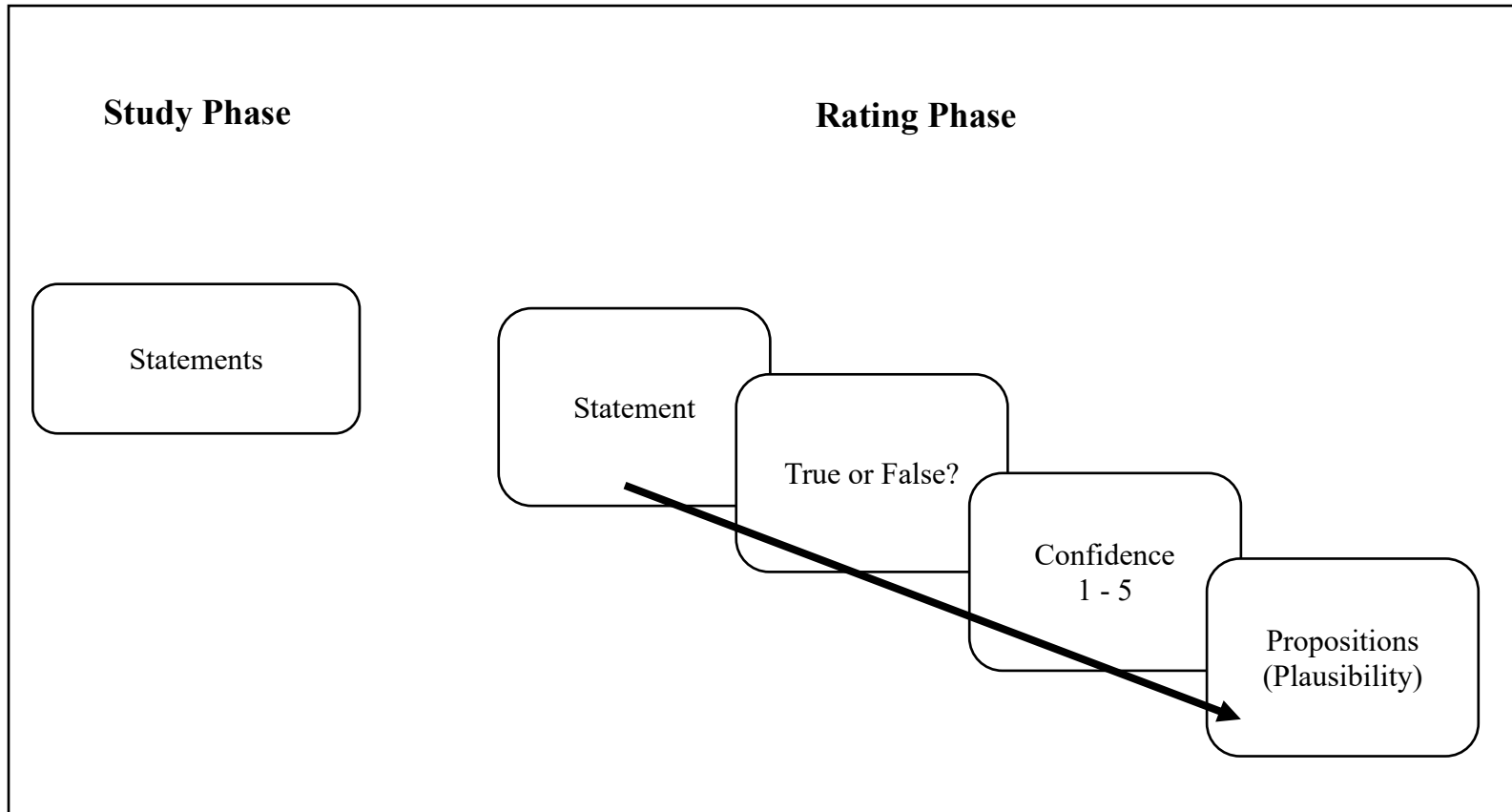


Figure 7. Diagram of Experiment 2 Procedure.

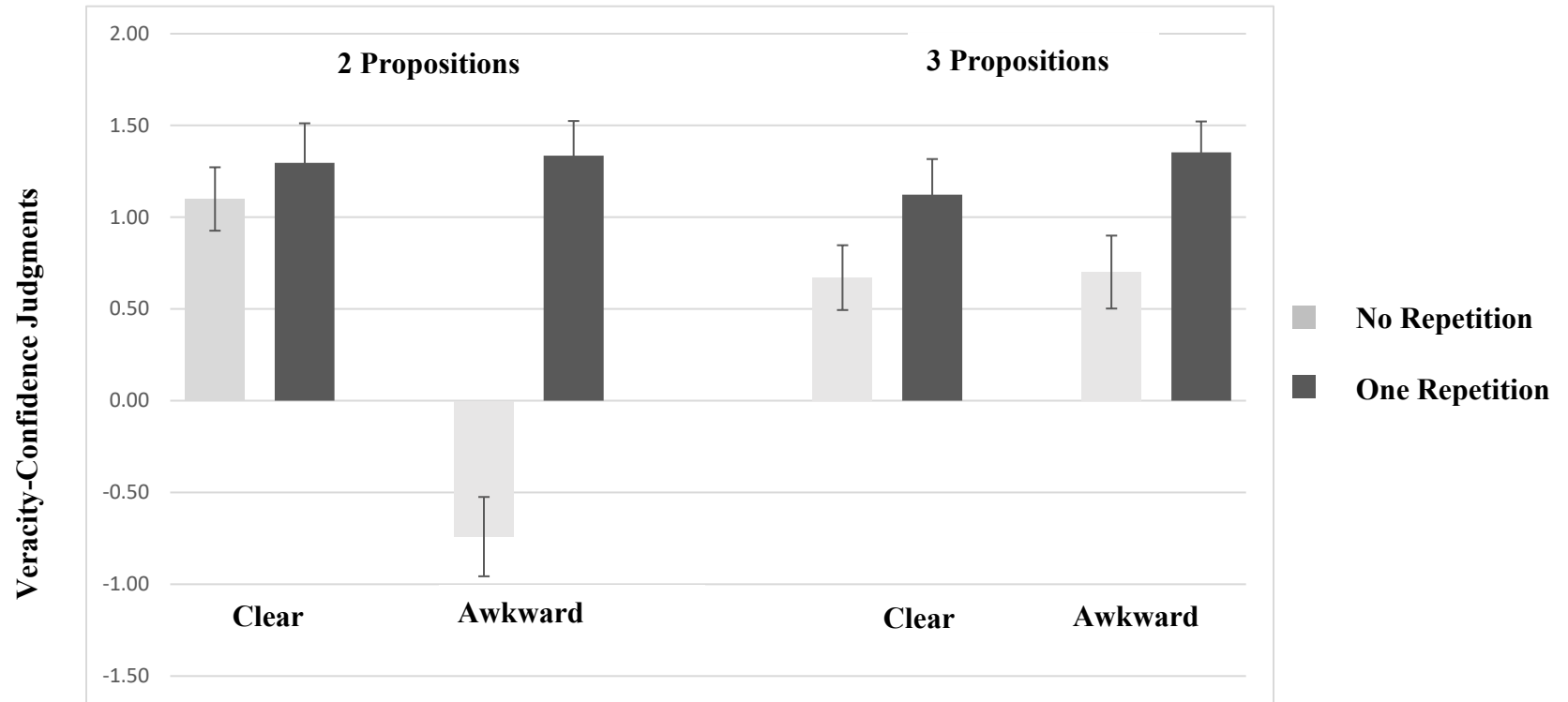


Figure 8. Experiment 2 ANOVA Results of Veracity Confidence Judgments by Repetition and Clarity for Statements with 2 and 3 Propositions.

Notes. The error bars indicate the standard error of the mean.

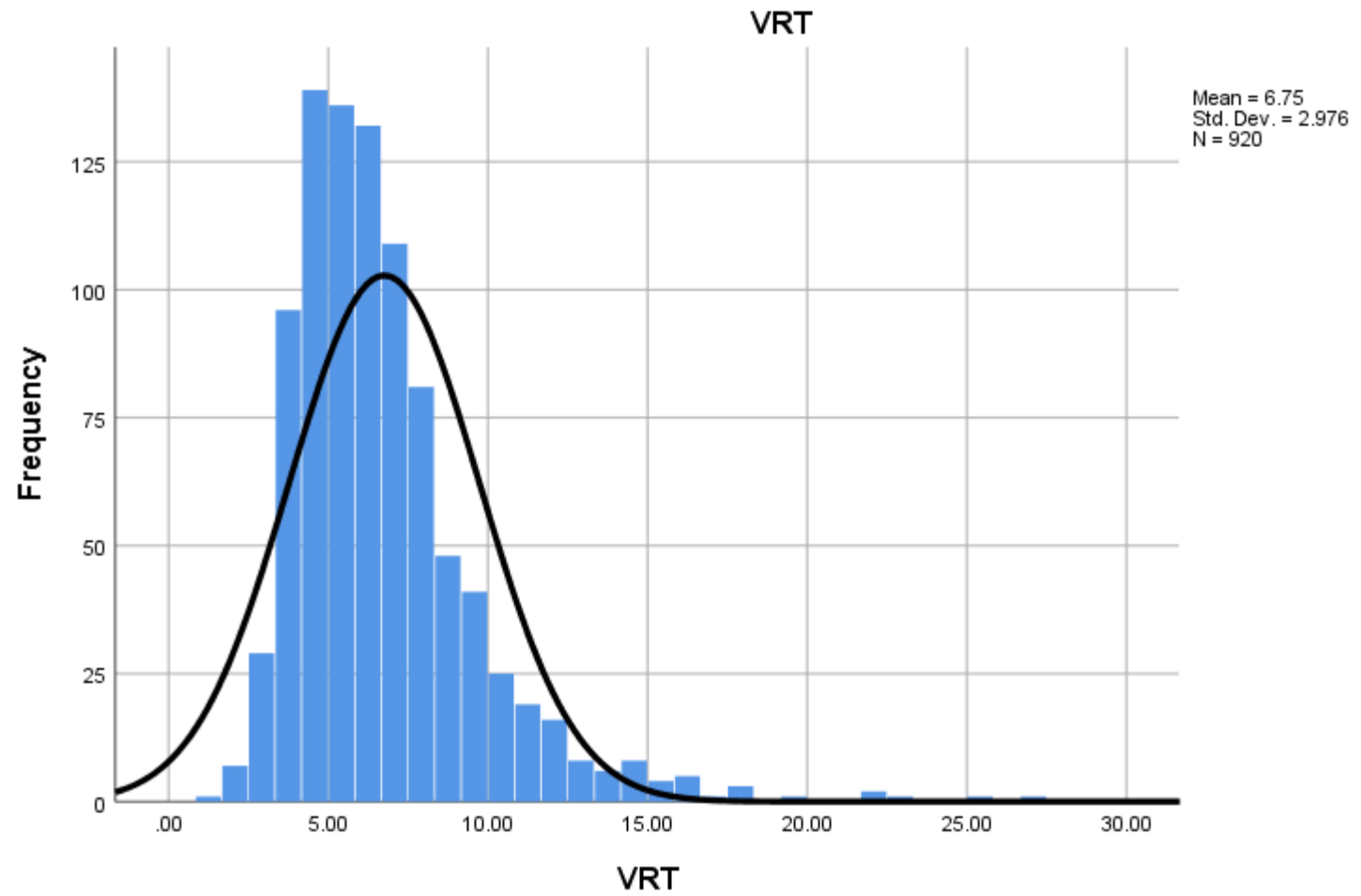


Figure 9. Histogram of Reaction Time Frequency for Experiment 2 Veracity Judgment Reaction Time.

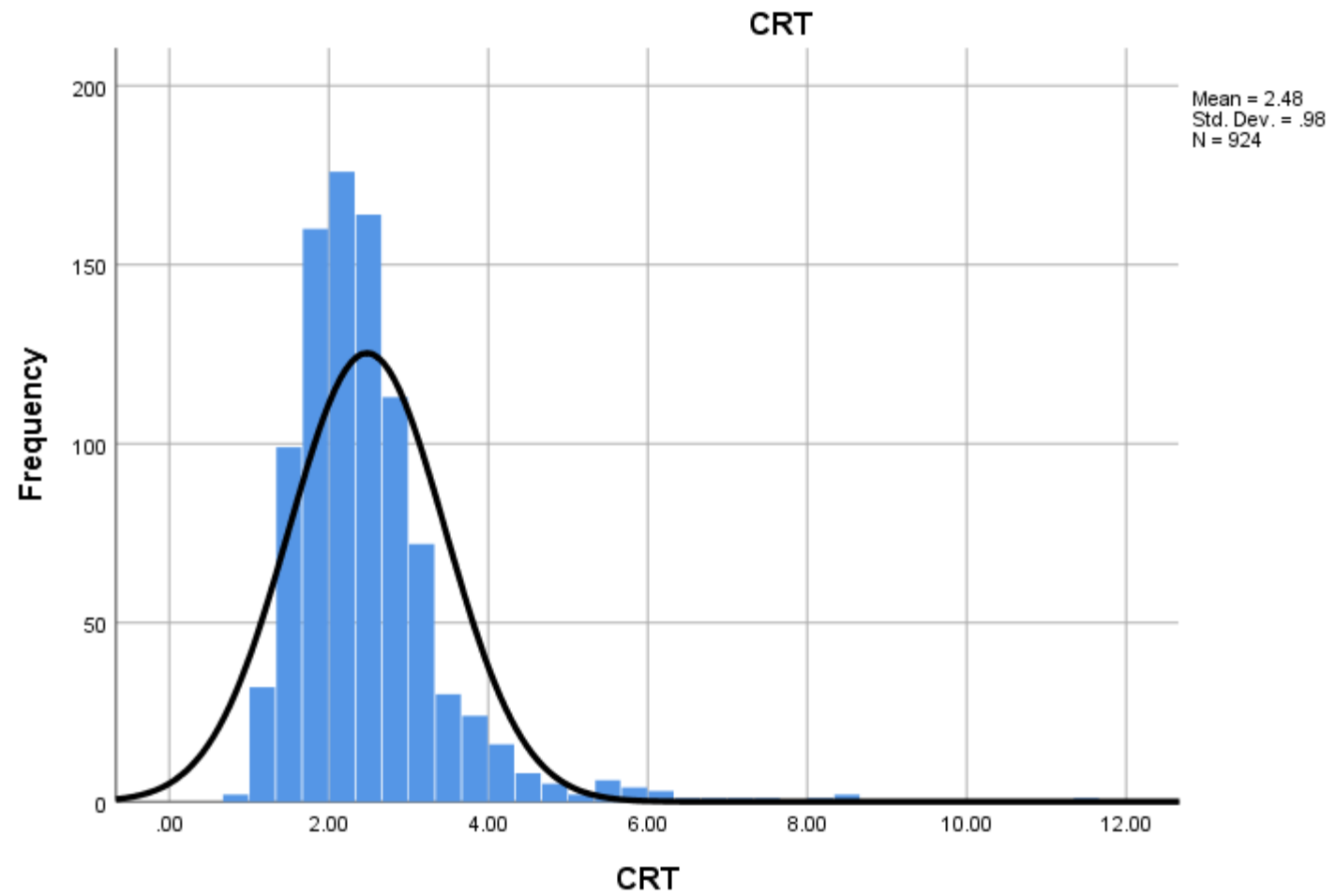


Figure 10. Histogram of Reaction Time Frequency for Experiment 2 Confidence Judgment Reaction Time.

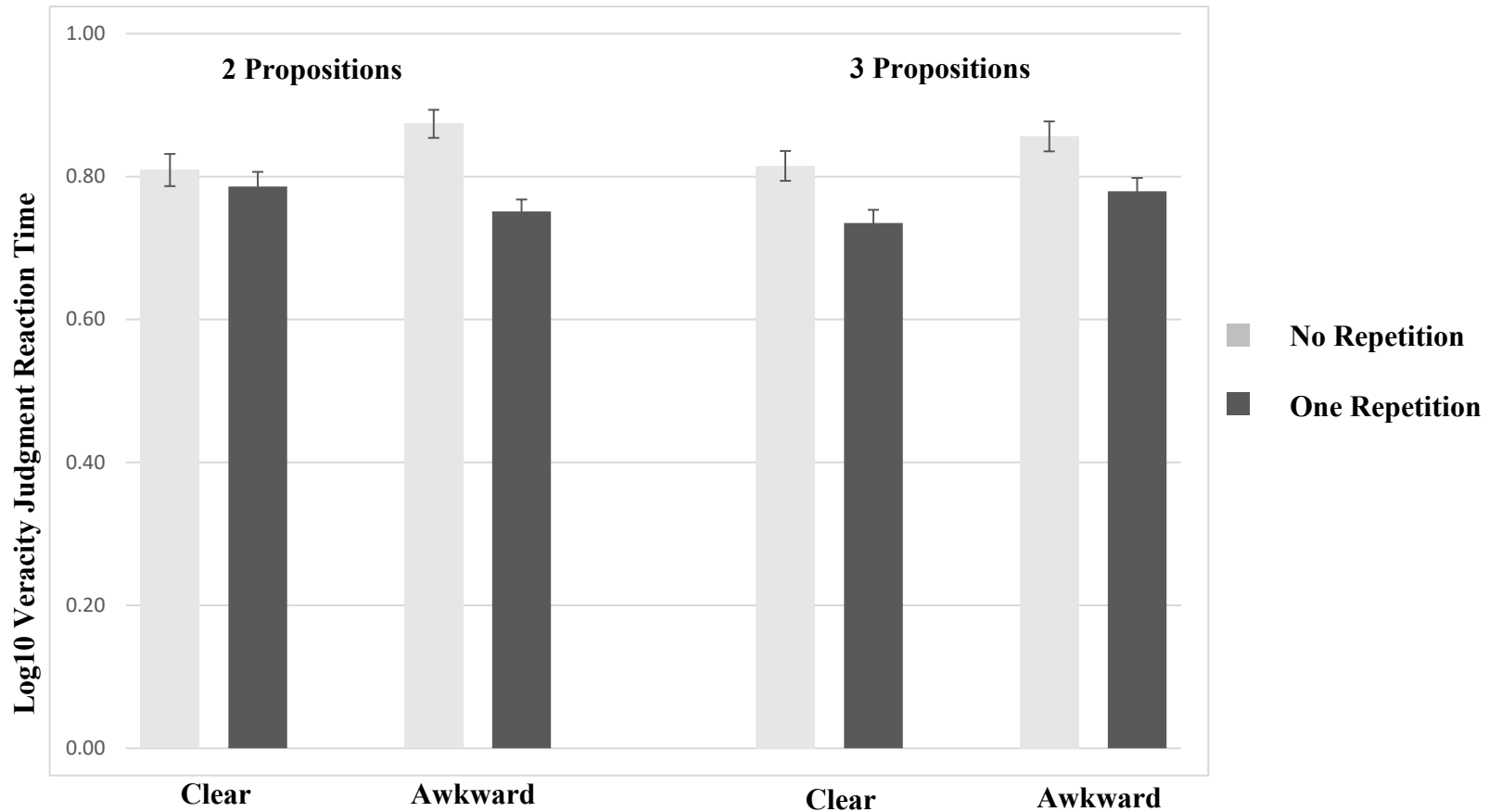


Figure 11. Experiment 2 ANOVA Results of Veracity Judgment Reaction Time by Repetition and Clarity for Statements with 2 and 3 Propositions.

Notes. The error bars indicate the standard error of the mean.

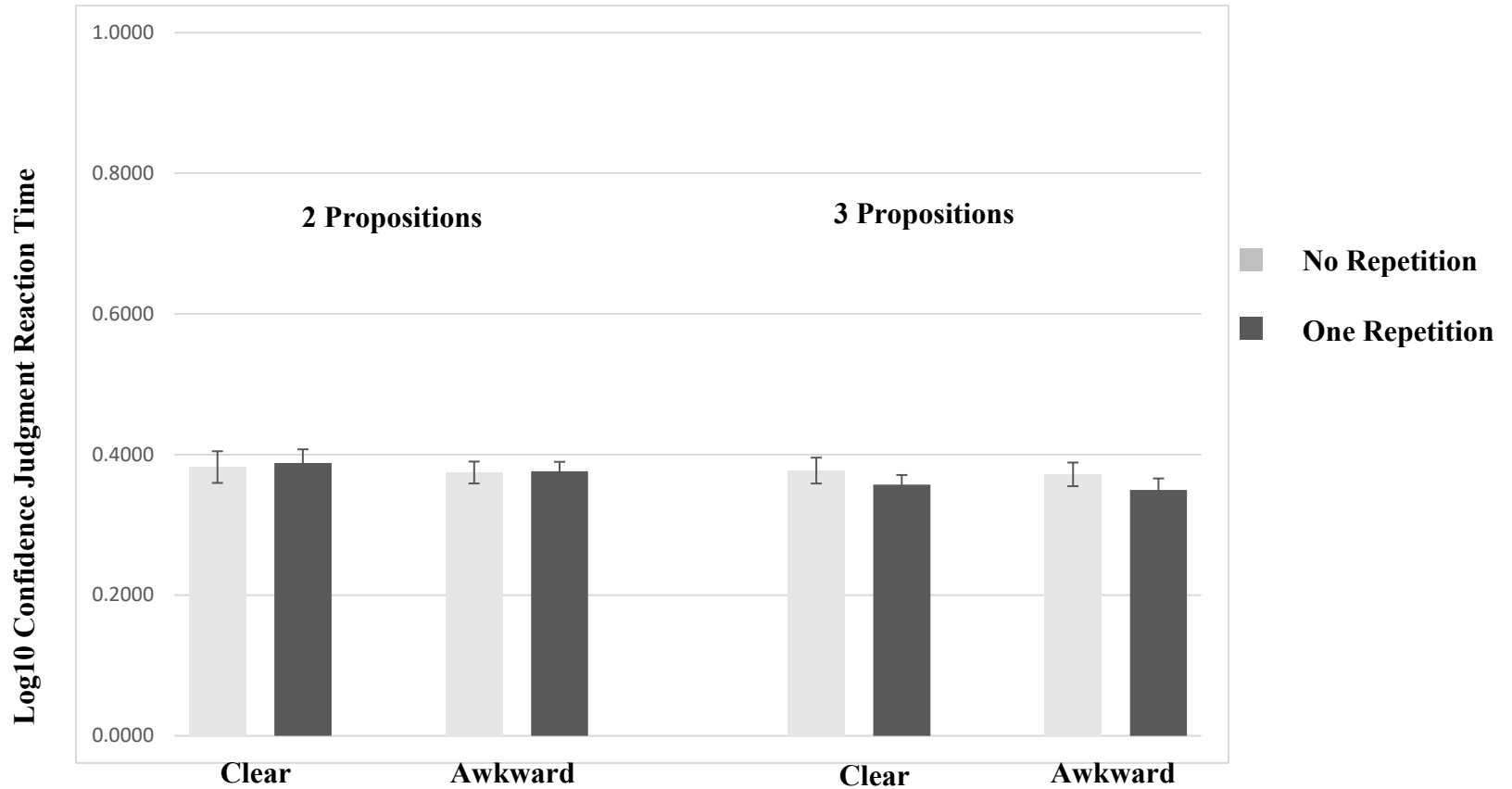


Figure 12. Experiment 2 ANOVA results of Confidence Judgment Reaction Time by Repetition and Clarity for Statements with 2 and 3 Propositions.

Notes. The error bars indicate the standard error of the mean.

APPENDIX B
PILOT STIMULI

TRUE	acc	conf	plau	False 1	acc	conf	plau	False 2	acc	conf	plau
Iron is atomic number 26.	61%	1	3	27	56%	1	3	25	39%	1	3
CNN began telecasting TV news in 1980.	83%	2	3	NBC	44%	2	3	Fox News	39%	1	3
Space Shuttle Challenger broke apart 73 seconds into its flight on January 28, 1986.	67%	1	3	Apollo	33%	1	3	Mercury	50%	1	2
An Ugli fruit is what you get if you cross a grapefruit, an orange and a tangerine.	22%	1	2	Kumquat	78%	1	3	Lychee	67%	3	3
Iron oxide gives Mars its reddish appearance.	89%	4	3	Lead tetroxide	28%	3	3	Copper Floride	56%	1	2
The Michelin Man is associated with tires.	94%	5	4	biscuits	83%	5	1	Candles	83%	5	4

The Woodstock festival was held at Max Yasgur's dairy farm during 3 days in 1969.	67%	1	3	Treasure Island	39%	1	2	Texxxas Jam	67%	1	2
Blue mountain coffee comes from Jamaica.	28%	1	3	Columbia	44%	1	2	Costa Rica	44%	1	3
The film Mamma Mia (2008) starred Pierce Brosnan.	83%	2	3	Collin Farrell	33%	1	3	Hugh Grant	33%	1	3
Stegosaurus means roof lizard.	44%	1	3	Triceratops	61%	1	2	Pterydactyl	72%	4	2
Lee Van Cleef played the role of 'the Bad' in the 1966 Spaghetti Western, The Good, the Bad and the Ugly.	56%	1	3	Clint Eastwood	67%	1	3	Eli Wallach	33%	1	3
The Spanish flag consists of red and yellow.	94%	5	3	Red & Green	44%	4	3	Green & Yellow	50%	5	2
The flag of Libya has the colors red, green, and black.	78%	3	3	Red, green, yellow	28%	1	3	Red, blue, black	28%	1	3

There are 4 C's that describe the properties of a diamond.	67%	3	3	3	22%	1	3	5	56%	1	3
Maman by Louise Bourgeois is a sculpture of a spider.	17%	1	2	giraffe	44%	1	1	Horse	72%	1	2
NASA stands for National Aeronautics and Space Administration.	78%	4	3	National Astronomical Science Association	33%	4	3	National Astronomy Science Administration	11%	4	4
For 9 years, future military strongman Idi Amin was Uganda's heavyweight boxing champion.	67%	1	3	Tanzania	33%	1	3	Rwanda	56%	1	3
The name of the lioness in the book Born Free is Elsa.	44%	1	2	Anna	50%	1	3	Emma	28%	1	2
10 Downing Street is the official address of Britain's prime minister.	50%	1	3	24 Sussex Ave.	67%	1	2	221-B Baker st.	33%	1	3

In a rainbow, red is at the outer part of the arch.	83%	5	4	violet	72%	5	2	orange	44%	2	4
Scurvy is a disease caused by a deficiency in Vitamin C.	72%	4	3	B	33%	1	2	D	72%	2	2
The hawksbill turtle belongs to the Cheloniidae family.	22%	1	2	Tuethida	72%	1	3	Nephropidae	72%	1	2
Matt Frewer portrayed the character Max Headroom.	50%	1	2	Harrison Ford	50%	1	2	Christopher Lambert	50%	2	3
Machu Picchu translates to old peak in English.	89%	2	3	Hidden temple	44%	1	3	Stone village	39%	1	3
Mickey Dolenz, Michael Nesmith, Peter Tork, and Davy Jones were members of the band The Monkees.	72%	1	3	The Animals	44%	1	1	The Buggles	39%	1	2

There are a total of 88 keys on a modern piano.	83%	1	3	87	56%	2	3	86	33%	2	3
Mel Blanc voiced the cartoon character Bugs Bunny.	83%	1	3	Don Messick	22%	1	3	Joseph Barbera	17%	1	3
There are 4 faces in a tetrahedon.	50%	4	3	3	72%	1	2	5	67%	5	3
The Vikings captured York in 866 AD.	50%	1	3	The Romans	61%	1	3	The Normans	44%	2	3
The nearest island country to Puerto Rico is the Dominican Republic.	89%	3	3	Trinidad	17%	1	3	Cuba	61%	3	3
Dom Perignon was the Benedictine monk who invented champagne.	11%	1	2	Moscatto	28%	1	3	Pinot-Griggio	61%	2	2
Lake Superior is the largest freshwater lake in the world.	67%	4	3	Great Salt Lake	61%	1	3	Lake Huron	44%	3	3
Mount Everest is the highest point on the	28%	1	2	K2	56%	1	3	Annapurna	11%	4	4

Earth's continental crust.											
Jack Ruby shot and killed Lee Harvey Oswald.	56%	1	2	Jack Opal	50%	2	3	Jack Jade	56%	1	2
Jolly Green Giant is the mirthful verdant colossus was created in 1926 by the Minnesota Valley Canning Company.	56%	1	2	Gulliver	67%	1	2	Paul Bunyan	33%	3	3
In Dutch folklore, Antigoon the giant was defeated by a young Roman soldier Brabo.	39%	1	2	Boeman	17%	1	2	Klaas Vaak	61%	1	3
The Greek word for diamond means unbreakable.	72%	2	3	priceless	39%	1	3	Precious	11%	2	3
Digitails is a plant commonly known as foxglove.	28%	1	2	penstemon	56%	1	2	Campanula	72%	1	3

Napoleon Bonaparte died in the battle of Trafalgar.	61%	2	2	Horatio Nelson	50%	1	3	Charles Martel	56%	1	3
The Pacific Ocean is the deepest ocean in the world.	44%	4	3	Atlantic Ocean	28%	2	3	Arctic Ocean	28%	2	3
Glinda is the name of one of the good witches in the Wizard of Oz.	94%	5	4	Gelinda	22%	5	4	Glenda	28%	1	3
Mercury is not a gas at room temperature.	39%	4	3	Hydrogen	83%	3	3	Helium	56%	4	3
The islands Hokkaido, Honshu, Shikoku and Kyushu are part of Japan.	56%	2	3	South Korea	33%	2	3	Vietnam	17%	3	3
In 2006, Pluto became reclassified as a dwarf planet.	61%	3	3	Mini planet	33%	4	3	Semi planet	17%	4	3
Aldous Huxley wrote Brave New World.	56%	1	3	Thomas Hardy	50%	1	3	D. H. Lawrence	33%	1	3
The largest contiguous land	83%	3	2	Roman Empire	22%	4	3	Persian Empire	33%	1	3

empire in history was the Mongol Empire.											
David Howell Evans is more widely known by his stage name: The Edge.	28%	1	3	The Cliff	61%	1	2	The Line	67%	1	3
As of 2016, there were 28 sports in the summer Olympics.	83%	1	3	34	28%	1	3	48	39%	2	3
Robert Baden-Powell was the founder of the Scout Movement.	44%	1	3	H. G. Wells	39%	1	3	Winston Churchill	61%	1	3
Ginger wine is made from ginger and raisins.	56%	2	3	chocolate	72%	1	2	Stout	39%	1	3
Lapsang Souchong is a type of tea.	39%	1	2	Type of language	78%	1	3	Breed of dog	56%	1	3
The Italian city Pisa has the famous leaning tower.	56%	5	3	Florence	33%	4	3	Rome	17%	4	3
Red tea has the least amount of caffeine among the types of tea.	33%	3	3	highest	61%	3	3	White tea	50%	1	3

The stone Blue John was found in Castleton, UK.	50%	1	3	Inverness, UK	72%	1	2	Kendal, UK	39%	1	3
Swaziland's government is an absolute monarchy.	50%	1	2	Communist oligarchy	50%	1	3	Absolute Theocracy	50%	1	2
In the film Forrest Gump, Forrest taught Elvis Presley a new way to dance.	39%	1	3	Fred Astaire	56%	4	3	Ginger Rogers	56%	1	3
Dry air roughly contains about 78% of nitrogen in volume.	50%	2	3	0.58	44%	3	3	0.38	28%	4	3
Nicaragua is the largest country in Central America.	44%	1	3	Guatemala	67%	2	3	Honduras	56%	3	3
Tchaikovsky composed The Nutcracker.	56%	3	3	Beethoven	50%	4	3	Mozart	44%	1	3
Simon De Bon is the lead singer of the band Duran Duran.	39%	1	3	Mar Almond	56%	1	3	Tony Hadley	50%	1	2

James Blunt's second album title is All the Lost Souls.	72%	1	3	Brave Souls	44%	1	3	Sad souls	28%	1	3
Capricorn is one of the earth signs.	72%	4	3	Leo	33%	5	4	Scorpio	44%	4	4
Tulips were once exchanged as a form of currency.	28%	3	2	Roses	72%	4	3	Lilies	61%	1	2
The winner of each stage of the Tour de France wears a yellow jersey in the next stage.	50%	1	3	green	33%	1	3	White	44%	1	3
The Cannes Film Festival was founded in 1946.	72%	1	3	1898	28%	1	3	1981	17%	1	3
As of 2018, there are 6 James Bond movies that have just one word in the title.	67%	2	3	5	56%	3	3	7	39%	1	3
Mozart composed The Marriage of Figaro.	39%	1	2	Puccinni	50%	2	2	Liszt	50%	1	3

A squirrel's nest is called a drey.	72%	1	3	eyrie	61%	1	3	sett	56%	1	2
Sports car manufacturer Ferrari is based in Maranello, Italy.	83%	3	3	Verona	11%	4	3	Ferrara	11%	3	3
The TGV (the French high-speed train) stands for Train a Grande Vitesse.	39%	1	3	Tactique Geographique Vierra	50%	1	3	Transport e Grandiose	28%	1	3
The medical term for lockjaw is tetanus.	39%	1	3	Rubella	72%	4	3	Mumps	100%	4	3
The poker player Gus Hansen's nickname is The Great Dane.	56%	1	3	Straightjack Gus	56%	1	2	The Devillish	39%	1	2
Graham's number is a power tower of 3.	39%	1	3	2	61%	1	2	5	50%	1	2
CIA stands for Central Intelligence Agency.	28%	3	3	Covert International Agency	17%	5	4	Centralized International Association	78%	4	3
The common name for Impatiens walleriana is busy Lizzie.	28%	1	2	Cuckooflower	83%	1	2	Foxglove	72%	1	2

Rice Krispies are represented by these characters: Crackle, Pop, and Snap.	33%	5	3	boom	11%	5	3	Snack	50%	5	4
India won the Hockey World Cup in 1975.	39%	1	3	Pakistan	72%	1	3	Germany	44%	1	3
The lion was lacking a heart in the Wizard of Oz.	89%	5	3	Tin man	44%	4	3	Scarecrow	61%	5	3
Lewis Wallace is the author of Ben-Hur.	39%	1	3	Bernard Shaw	61%	1	2	Victor Hugo	50%	1	3
Quercus is the Latin name of an oak tree.	72%	1	3	Greek coin	61%	1	3	Red squirrel	56%	1	3
The shamrock is the national flower of Ireland.	89%	1	3	The marigold	28%	4	3	The daffodil	17%	2	3
Switzerland is a landlocked country.	39%	1	2	Spain	28%	1	3	France	72%	1	3
The Hope Diamond is the largest blue diamond in the world at 45 carats and is said to be cursed.	72%	1	2	The star of Africa	39%	1	2	Koh-l-noor	44%	1	3

The birthplace of the fictional detective Hercule Poirot is Belgium.	39%	1	2	France	39%	1	2	England	61%	1	2
More than 10% of the world's Muslims live in Indonesia.	72%	1	3	Iran	67%	1	3	Saudi Arabia	33%	2	3
A Rhode Island Red is a breed of chicken.	39%	1	3	duck	72%	1	2	Pig	78%	1	3
In Star Trek, Captain Kirk's middle name is Tiberius	67%	1	3	Thaddeus	39%	1	3	Tadleigh	39%	1	3
Cuneiform script is the earliest known writing system in the world.	50%	1	3	Prakirt script	44%	3	3	Sanskript	39%	1	3
In the 1971 film Dirty Harry, Clint Eastwood's character's last name is Callahan.	50%	1	3	Colt	28%	1	3	Carnegie	50%	1	3
Gold is the chemical element that has the atomic number of 79.	44%	2	3	Silver	44%	1	3	Nickel	39%	1	3

Brazil is the only Portugese speaking country in the Americas.	22%	1	3	Argentina	50%	4	3	Colombia	78%	3	3
Wellington is the capital of New Zealand.	72%	1	3	Auckland	50%	1	2	Christchurch	61%	1	3
Lake Allatoona is in the state of Georgia.	56%	1	3	Alabama	67%	1	2	Colorado	39%	1	3
The capital city of Estonia is Tallinn.	61%	1	2	Vilnius	50%	1	2	Riga	39%	1	3
The musical instrument trumpet has three valves.	72%	1	3	four	28%	2	3	Two	67%	4	3
Catherine of Aragon was Henry VIII's first wife.	56%	1	3	second	39%	1	3	third	22%	1	3

APPENDIX C

EXPERIMENT 1 STIMULI

Target	Confirming	Neutral	Contradicting
<p>Iron is atomic number 26.</p> <p>(61% acc, conf 1, plau 3)</p>	<p>The atomic number is determined by the number of protons.</p> <p>Iron has 26 protons.</p>	<p>Iron can be found in your blood stream.</p> <p>Dark leafy greens are good sources of iron.</p>	<p>Iron is a metal in the first transition series.</p> <p>The mass number of Iron is 55.845.</p>
<p>CNN began telecasting TV news in 1980.</p> <p>(83% acc, conf 2, plau 3)</p>	<p>CNN was founded by American media proprietor Ted Turner.</p> <p>First television channel to provide 24 hour coverage was CNN.</p>	<p>The latest news can be found on CNN.</p> <p>CNN is popular news broadcasting station.</p>	<p>CNN was founded after NBC.</p> <p>CNN began telecasting TV sports news in 1980.</p>
<p>Space Shuttle Challenger broke apart 73 seconds into its flight on January 28, 1986.</p> <p>(67% acc, conf 1, plau 3)</p>	<p>Nasa knew that the design of the Space Shuttle Challenger could fail prior to the event.</p> <p>Ronald Reagan issued a special investigation into the Space Shuttle Challenger accident.</p>	<p>The Space Shuttle Challenger was going to orbit around the earth.</p> <p>There is a mvie about the Space Shuttle Challenger.</p>	<p>The flight of the Space Shuttle Challenger was successful for 9 missions.</p> <p>The Apollo 13 mission has been dramatized multiple times.</p>

<p>The Woodstock festival was held at Max Yasgur's dairy farm during 3 days in 1969.</p> <p>(67% acc, conf 1, plau 3)</p>	<p>Many people gathered for the first Woodstock Festival.</p> <p>the Woodstock Festival first began in 1969</p>	<p>Many artists perform at the Woodstock Festival.</p> <p>The Woodstock festival will celebrate it's 50th anniversary this year.</p>	<p>The Woodstock festival was referred to as 3 days of Peace & Music.</p> <p>The Woodstock Festival of 1969 was held at a 600-acre dairy in Bethel, New York.</p>
<p>Blue mountain coffee comes from Jamaica.</p> <p>(28% acc, conf 1, plau 3)</p>	<p>The Blue mountains is near Kingston where some of the highest mountains in the Caribbean are located. Jamaica is known for their gourmet products.</p>	<p>Blue Mountain coffee comes in different roasts.</p> <p>Blue Mountain coffee is very expensive.</p>	<p>The coffee plants are originated in Ethiopia.</p> <p>Jamaica grows just 0.1% of the world's coffee.</p>
<p>The film Mamma Mia (2008) starred Pierce Brosnan.</p> <p>(83% acc, conf 2, plau 3)</p>	<p>Pierce Brosnan stepped out of his usual role when he joined the musical film, Mamma Mia (2008). Pierce Brosnan has won multiple acting related rewards and was famous for his role of James Bond.</p>	<p>The film Mamma Mia (2008) is a romantic musical.</p> <p>The film Mamma Mia (2008) was made into a sequel.</p>	<p>Pierce Brosnan is famous for the James Bond role.</p> <p>The American singer, Cher, starred in the movie Mamma Mia (2008), and played as Ruby Sheridan.</p>
<p>Lee Van Cleef played the role of 'the Bad' in the 1966 Spaghetti Western, The Good, the Bad and the Ugly.</p> <p>(56% acc, conf 1, plau 3)</p>	<p>Lee can Cleef starred in multiple Western movies throughout his career.</p> <p>Lee Van Cleef achieved stardom in 1965 which lasted 12 years.</p>	<p>"If you work for a living, why do you kill yourself working?" is a popular quote from the movie "the Good, the Bad, and the Ugly."</p> <p>Tuco the Ugly was a main character in the film The good, the Bad and the Ugly.</p>	<p>Lee Van Cleef was best known for playing in the spaghetti western, For a Few Dollars More.</p> <p>Clint Eastwood played the role of "the Good" in The Good, the Bad, and the Ugly.</p>

<p>The flag of Libya has the colors red, green, and black.</p> <p>(78% acc, conf 3, plau 3)</p>	<p>The Pan African flag is red, green, and black.</p> <p>Lybia was an early center of Christianity.</p>	<p>The flag Libya has a simple design.</p> <p>Libya's flag symbolizes independence.</p>	<p>The flag of Libya contains a white crescent moon and star.</p> <p>The Flag of the Great Socialist People's Libyan Arab Jamahiriya was a plain green flag.</p>
<p>Maman by Louise Bourgeois is a sculpture of a spider.</p> <p>(17% acc, conf 1, plau 2)</p>	<p>Maman by Louise Bourgeois was and ode to her mother.</p> <p>The spider is supposed to be a symbol of strength and protection.</p>	<p>Louise Bourgeois is well-known for large scale sculptures.</p> <p>Louise Bourgeois began her life as an artist in Paris, France.</p>	<p>Louise Bourgeois was a painter and printmaker.</p> <p>Plenty of Louise Bourgeois's work displayed themes of sexuality and body.</p>
<p>For 9 years, future military strongman Idi Amin was Uganda's heavyweight boxing champion.</p> <p>(67% acc, conf 1, plau 3)</p>	<p>Idi Amin, the former President of Uganda was described as very tall and well built.</p> <p>His first championship was as the lightweight boxing champion.</p>	<p>Idi Amin began boxing when he was serving in the military.</p> <p>Idi Amin was a highly ranked soldier in the military.</p>	<p>Idi Admin was the president of Uganda and was notorious for his oppressiveness.</p> <p>Idi Amin left school with only a fourth grade English language education.</p>
<p>The name of the lioness in the book Born Free is Elsa.</p> <p>(44% acc, conf 1, plau 2)</p>	<p>Born free was also made into a British drama film in 1966 starting Virginia McKenna.</p> <p>Joy and George Adamson, a real-life couple raised the lioness.</p>	<p>Born Free is a book about the experiences of raising a lion cub.</p> <p>The book Born Free was made into a movie in the 1960s.</p>	<p>The book born free is based on a true story in Kenya.</p> <p>The real name of the lion that starred in the movie Born Free is Ugas.</p>

<p>10 Downing Street is the official address of Britain's prime minister.</p> <p>(50% acc, conf 1, plau 3)</p>	<p>Downing Street has traditionally been occupied by Britain's prime ministers. The address 10 Downing Street is nearly 300 years old and holds much significance.</p>	<p>Britain's current prime minister is Theresa May.</p> <p>Theresa May has been Britain's prime minister since 2016.</p>	<p>12 Downing Street is the official address of Britain's prime minister.</p> <p>10 Downing Street is the official address of Jamaica's prime minister.</p>
<p>The hawksbill turtle belongs to the Cheloniidae family.</p> <p>(22% acc, conf 1, plau 2)</p>	<p>Hawksbill turtles are extremely endangered.</p> <p>Sea turtles in the Cheloniidae family migrate thousands of miles a year in order to breed.</p>	<p>The hawksbill turtle is critically endangered.</p> <p>Hawksbill turtles live in rocky and shallow coastal areas.</p>	<p>Cheloniidae family has round shells and paddle like flippers.</p> <p>There are six species of turtles that belong to the Cheloniidae family.</p>
<p>Matt Frewer portrayed the character Max Headroom.</p> <p>(50% acc, conf 1, plau 2)</p>	<p>Matt Frewer has played pivotal roles in dozens of films throughout his career. Matt Frewer graduated acting school in the early 1980's.</p>	<p>Max Headroom is a fictional artificial intelligence character.</p> <p>The Max Headroom Show was a television series in the UK.</p>	<p>Matt Ford portrayed the character Max Headroom</p> <p>Matt Ford portrayed the dog Max Headroom</p>
<p>Machu Picchu translates to old peak in English.</p> <p>(89% acc, conf 2, plau 3)</p>	<p>Machu Piccu is located on a mountain ridge that is 2,430 metres high. Often mistakenly referred to as "Lost City of the Incas", Machu Picchu was brought to attention by an American historian in 1911.</p>	<p>Machu Picchu is a citadel in the Andes Mountains.</p> <p>The former use of Machu Picchu is a mystery.</p>	<p>Machu Picchu translates to big peak in English</p> <p>Machu Picchu was abandoned during the Spanish Conquest.</p>
<p>Mickey Dolenz, Michael Nesmith, Peter Tork, and Davy Jones were</p>	<p>The Monkees were a band with four members.</p>	<p>Davy Jones was the lead singer of his band, backed up</p>	<p>One of the members of the band Monkees told a lie to</p>

<p>members of the band The Monkees.</p> <p>(72% acc, conf 1, plau 3)</p>	<p>The Monkees were a pop rock band.</p>	<p>by Mickey Dolenz, Michael Nesmith, and Peter Tork. Mickey Dolenz, Michael Nesmith, Peter Tork, and Davy Jones were all members of a popular rock band.</p>	<p>the reporter that they sold more than the Beatles.</p> <p>There was more than one drummer in the band the Monkees.</p>
<p>There are a total of 88 keys on a modern piano.</p> <p>(83% acc, conf 1, plau 3)</p>	<p>The piano has an even number of keys total.</p> <p>There are black and white keys on a piano.</p>	<p>It takes a lot of skill to play the modern piano.</p> <p>Deep bass and high treble notes can be played on the modern piano.</p>	<p>There total number of keys on a modern piano is an odd number.</p> <p>The Imperial Concert Grand Bosendorfer piano has 97 keys.</p>
<p>Mel Blanc voiced the cartoon character Bugs Bunny.</p> <p>(83% acc, conf 1, plau 3)</p>	<p>Mel Blanc was often referred to as the "Man of a Thousand Voices"</p> <p>Mel Blanc voiced dozens of classic cartoon characters.</p>	<p>Bugs Bunny is famous character in the show Looney Toons.</p> <p>Bugs Bunny is characterized by his Brooklyn accent and most known for his catch phrase, "eh, what's up doc?"</p>	<p>Bugs bunny was voiced by Greg Burson in the 1990s.</p> <p>Mel Blanc was well known for his role as Barney Rubble of the Flintstones.</p>
<p>The Vikings captured York in 866 AD.</p> <p>(50% acc, conf 1, plau 3)</p>	<p>Vikings tended to raid European areas quite regularly</p> <p>The viking age dated back to 790 A.D- 1066 A.D.</p>	<p>York is a historic walled city in North Yorkshire, England.</p> <p>York is most known for its 13th-century gothic cathedral, York Minister.</p>	<p>Two Vikings ascended to the throne of England 1013-1035 AD.</p> <p>The Viking Age may be expanded to North Germanic dominance including Scandinavian York.</p>
<p>The nearest island country to Puerto Rico is the Dominican Republic.</p>	<p>Both Puerto Rico and the Dominican Republic are part of a group of Island</p>	<p>Puerto Rico is a US territory and is a popular tourist destination.</p>	<p>The furthest island country to Puerto Rico is the Dominican Republic.</p>

(89% acc, conf 3, plau 3)	countries in the Caribbean Sea. The Dominican Republic is the smallest island of the Greater Antilles island chain.	San Juan is the capital of Puerto Rica and was founded in 1521.	Cuba is the nearest island country to Puerto Rico.
Dom Perignon was the Benedictine monk who invented champagne. (11% acc, conf 1, plau 2)	Dom Perignon is a popular brand of champagne. Champagne is known to have been accidentally created by a monk in 1600s	Champagne was unintentionally invented in 1693. Champagne is a sparkling wine and is often served at celebratory events.	Benedictine Monks observed abstinence. Christopher Merret was the first to document wine with sugar added to be "sparkling wine".
An Ugli fruit is what you get if you cross a grapefruit, an orange, and a tangerine. (22% acc, conf 1, plau 2)	An Ugli fruit looks like an uglier version of an orange which is why it got its name. An Ugli fruit is known to be less bitter than a grapefruit.	The highest point on Earth's continental crust is almost 30,000 feet above sea level. Several adventurers attempt to climb the highest point on Earth every year.	An Ugli fruit is often guessed to be a lemon-tangerine hybrid. A pomelo is often used for an Ugli fruit hybridization.
Jack Ruby shot and killed Lee Harvey Oswald. (56% acc, conf 1, plau 2)	Jack Ruby was jailed on March 14, 1964. Ruby died in Jail after his murder conviction.	Lee Harvey Oswald was shot and killed in 1963. Lee Harvey Oswald assassinated the 35th US president.	Lee Harvey Oswald assassinated John F. Kennedy. Jack Ruby was a Dallas Texas Nightclub owner at the time of JFK's assassination.
Jolly Green Giant is the mirthful verdant colossus was created in 1926 by	The jolly green giant was created in response to the discovery of a new type of	The Minnesota Valley Canning Company was founded in 1903 in Le Sueur, Minnesota.	Jolly Green Giant is the dejected verdant colossus was created in 1926 by the

<p>the Minnesota Valley Canning Company.</p> <p>(56% acc, conf 1, plau 2)</p>	<p>pea which was especially large.</p> <p>Minnesota Valley Canning Company was popular for its canned vegetables.</p>	<p>Minnesota Valley Canning Company was a mass producer of canned corn and peas.</p>	<p>Minnesota Valley Canning Company.</p> <p>The Minnesota Valley Canning Company was renamed to the Green Giant Company in 1950.</p>
<p>In Dutch folklore, Antigoon the giant was defeated by a young Roman soldier Brabo.</p> <p>(39% acc, conf 1, plau 2)</p>	<p>This is an origin story for the city of Antwerp and that is where the statue is located.</p> <p>A statue is of Brabo holding up Antigoon's hand which he cut off.</p>	<p>There is a monument dedicated to the Roman soldier Brabo in Belgium.</p> <p>Brabo is a character in Dutch folklore that is known for defeating the giant.</p>	<p>In Greek folklore, Antigoon the giant was defeated by a young Roman soldier Brabo.</p> <p>The Antwerp fountain is at Antwerp where the mythical giant Antigoon lived.</p>
<p>The Greek word for diamond means unbreakable.</p> <p>(72% acc, conf 2, plau 3)</p>	<p>Diamonds are a notoriously strong gem.</p> <p>Diamonds were rarely used pre 1300s due to the material being very difficult to work with.</p>	<p>Diamonds are the birth stone for the month of April.</p> <p>Diamonds have become known to symbolize love and fidelity.</p>	<p>Natural diamonds and imitation diamonds can be distinguished using optical techniques.</p> <p>The element carbon with its atoms arranged in a crystal structure is called diamond cubic.</p>
<p>Digitails is a plant commonly known as foxglove.</p> <p>(28% acc, conf 1, plau 2)</p>	<p>The flower Digitalis is shaped like a finger and actually covers up a person's finger.</p> <p>Digitalis in Latin translates to "finger-like"</p>	<p>Digitalis can be used to treat heart failure by increasing blood flow.</p> <p>Digitalis can be found in a wide range of colors.</p>	<p>The genus of the plant Digitalis is placed in the figwort family.</p> <p>The colors of the Digitalis is various from pink to light grey.</p>
<p>Mount Everest is the highest point on the Earth's continental crust.</p>	<p>The summit of mount Everest has a larger</p>	<p>The Battle of Trafalgar took place on October 21, 1805.</p>	<p>K2 is known as the Savage Mountain as it is extremely hard to climb.</p>

(28% acc, conf 1, plau 2)	<p>atmospheric pressure than any other mountain</p> <p>Mount Everest is one of the most dangerous mountains to climb.</p>	<p>The battle of Trafalgar was fought by the British Royal Navy against the combined fleets of the French and Spanish Navies.</p>	<p>Mountains near the equator are technically higher than those in other areas.</p>
<p>The islands Hokkaido, Honshu, Shikoku and Kyushu are part of Japan.</p> <p>(56% acc, conf 2, plau 3)</p>	<p>Japan is a made up of multiple large islands.</p> <p>Japan is not located on a single landmass.</p>	<p>Hokkaido is known for its volcanoes and natural hot springs.</p> <p>Shikoku is an island home to 8 of the pilgrimage temples.</p>	<p>About 45km north to Hokkaido lies Sakhalin Island, Russia.</p> <p>The Chinese characters for Shikoku literally means 4 provinces and Kyushu means 9 islands.</p>
<p>Aldous Huxley wrote Brave New World.</p> <p>(56% acc, conf 1, plau 3)</p>	<p>Aldous Huxley was a famous writer and philosopher.</p> <p>Aldous Huxley was the author of nearly 50 books.</p>	<p>Brave New World is a dystopian novel published in 1932.</p> <p>The book Brave New World was made into a movie in 1998.</p>	<p>Aldous Huxley was a character in Brave New World.</p> <p>Aldous Huxley directed Brave New World.</p>
<p>David Howell Evans is more widely known by his stage name: The Cliff.</p> <p>(61% acc, conf 1, plau 2)</p>	<p>David Howell Evans is a famous rock guitarist</p> <p>David Howell Evan's nickname derives from the shape of his head.</p>	<p>David Howell Evans is an Irish musician and songwriter.</p> <p>David Howell Evans has recorded 14 studio albums with the band U2.</p>	<p>David Howell Evans is more widely known by his stage name: The Edge.</p> <p>David Howell Evans is most known for being member of the band U2.</p>
<p>As of 2016, there were 34 sports in the summer Olympics.</p>	<p>There were a number of additions to Olympic sports in 2016.</p>	<p>The summer Olympics is an event held every four years.</p>	<p>There are an odd number of sports in the summer Olympics in honor of the</p>

(28% acc, conf 1, plau 3)	Olympic sports haven't been removed in decades.	The next summer Olympics will take place on July 24, 2020.	odd numbered Olympic rings. Only five sports have been contested at every summer Olympic Games since 1896.
Winston Churchill was the founder of the Scout Movement. (61% acc, conf 1, plau 3)	The scout movement was founded in the 1940s. The scout movement was created in response to WWII.	The Scout Movement has over 50 million participants. The Scout Movement was founded in 1922 in Paris, France.	Lord Baden-Powell was the founder of the Scout Movement. Winston Churchill led Britain through World War II.
Ginger wine is made from ginger and chocolate. (72% acc, conf 1, plau 2)	Ginger wine Is known to have two key ingredients. The fat in the chocolate and water in alcohol repel each other.	Ginger wine was first produced in England. Ginger wine is believed to help relieve flu symptoms.	Ginger wine is made from ginger and raisins. Sales of Ginger Wine boosted in the 19th Century due to a Cholera epidemic.
Lapsang Souchong is a breed of dog. (56% acc, conf 1, plau 3)	Laspang Souchong originated from mountainous regions of china as guide dogs. Laspang Souchong is a popular breed of dog in Asia.	Lapsang Souchong originates from Fujian, China. Dog breeds are sometimes named after a location.	Lapsang Souchong is a tea well-known as the "tea for Westerners". Lapsang Souchong is not on the American Kellen Club's official list of dog breeds.
The stone Blue John was found in Inverness, UK. (72% acc, conf 1, plau 2)	Blue John is a type of rare stone. Blue John are only found in the UK	Blue John is a semi-precious mineral that is only found in Blue John Cavern. The stone Blue John is believed to be first discovered by the Romans.	Inverness, UK is the largest city of the Scottish Highlands that used to have great food markets. The stone Green John is found only in Green Valley.

Swaziland's government is an absolute theocracy. (50% acc, conf 1, plau 2)	Swaziland keeps traditions a core part of their political system. Many countries today still use theocracy as a form of their government.	Swaziland has two capitals, Mbabane and Lobamba. Swaziland is located in southern Africa and is known for it's wilderness reserves.	Afghanistan's government is an absolute theocracy. The Swaziland's current king has 15 wives.
In the film Forrest Gump, Forrest taught Ginger Rogers a new way to dance. (56% acc, conf 1, plau 3)	Forrest Gump is known to be a great dancer. One of the most popular scenes in the film is when Forrest teaches a boy how to dance.	In the film Forrest Gump, Forrest is known for his famous line "life is like a box of chocolates, you never know what you're gonna get." In the film Forrest Gump, Forrest received a Medal of Honor after fighting in the Vietnam War	In the film Forrest Gump, we see Elvis Presely dance. In the film Forrest Gump, Jenny taught Forrest a new way to dance.
Tony Hadley is the lead singer of the band Duran Duran. (50% acc, conf 1, plau 2)	Duran Duran is a famous band that formed in the 1970s. Duran Duran have 4 members.	Duran Duran was a band formed in Birmingham in 1978. "A View to the Kill" was a popular song by the band Duran Duran.	Simon Le Bon is the lead singer in the band Duran Duran. Tony Hadley is the lead singer of the band Spandau Ballet.
James Blunt's second album title is All the Brave Souls. (44% acc, conf 1, plau 3)	James Blunt is a singer and songwriter known for his ballads. James Blunt is most well known for his song "You're beautiful".	James Blunt's real name is James Hillier Blount and he rose to fame in 2004. James Blunt is an English singer-songwriter, record producer and former British Army Officer	James Blunt's second album title is All the Lost Souls. James Blunt's third album was titled Some Kind of Trouble.
The winner of each stage of the Tour de France	The Tour de France implemented rules to help	The Tour de France is an annual men's bicycle race.	The winner of each stage of the Tour de France wears a

<p>wears a white jersey in the next stage.</p> <p>(50% acc, conf 1, plau 3)</p>	<p>keep track of the current lead for each stage after having a hard time keeping track.</p> <p>Special jerseys are awarded to individuals who display extraordinary performance in the Tour de France.</p>	<p>The winner of each stage in the Tour de France wears a different colored jersey in the following stages.</p>	<p>yellow jersey in the next stage.</p> <p>In one of the classifications of the Tour de France, the winner wears a white jersey with red polka dots.</p>
<p>The Cannes Film Festival was founded in 1898.</p> <p>(28% acc, conf 1, plau 3)</p>	<p>This festival was held just 10 years after the first motion film</p> <p>The first kinetoscope film shown publicly was made in 1893.</p>	<p>The Cannes Film Festival is held in France each year which previews new films of all genres.</p> <p>The Cannes Film Festival was originally called the International film festival.</p>	<p>The Cannes Film Festival was founded in 1946.</p> <p>More films were made in color in the Mid 1950's.</p>
<p>Liszt composed The Marriage of Figaro</p> <p>(50% acc, conf 1, plau 3)</p>	<p>The opera Marriage of Figaro is written in Italian while the composer was not Italian.</p> <p>Liszt was a Hungarian artist of the Romantic Era.</p>	<p>The Marriage of Figaro is an opera buffa and was written in 1786.</p> <p>The Marriage of Figaro is an opera that takes place in Spain in the 18th century.</p>	<p>Franz Liszt was a Hungarian composer.</p> <p>Wolfgang Amadeus Mozart composed the Marriage of Figaro.</p>
<p>A squirrel's nest is called a sett.</p> <p>(56% acc, conf 1, plau 2)</p>	<p>A squirrel's nest is named as such, because squirrels never relocate their nests.</p> <p>This name was earned over 200 years ago</p>	<p>Squirrels build nests out of twigs, dry leaves and grass.</p> <p>Squirrels build nests to keep warm in the winter months.</p>	<p>A squirrel's nest is called a drey.</p> <p>A badger's nest is called a sett.</p>
<p>The TGV (the French high-speed train) stands for Tactique Geographique Vierra.</p>	<p>The TGV (the French high-speed train) has had many names in the past.</p>	<p>The TGV is France's intercity high-speed rail service, operated by the</p>	<p>The TGV extends its service to Belgium, Germany, and even the Netherlands.</p>

(55% acc, conf 1, plau 3)	The TGV started off with intercity transportation but expanded to international transportation.	SNCF, the state-owned national rail operator. The TGV is a high-speed train that began operating in 1981.	The abbreviation TGV stands for a French saying that translates to high-speed train.
The poker player Gus Hansen's nickname is Straightjack Gus. (56% acc, conf 1, plau 2)	Gus Hansen was notorious for pulling straight's in tournaments. Earned this nickname after making a series of risky bluffs in his most famous tournament appearance.	Gus Hansen is a famous Danish poker player who lives in Monaco. Gus Hansen has won three World Poker Tour open titles.	Gus Hansen was already a world class backgammon player. Hansen regularly plays in the "Bobby's Room" at the Bellagio Casino in Las Vegas.
Graham's number is a power tower of 2. (61% acc, conf 1, plau 2)	A power tower often referred to as a tetration, is the next hyperoperation after exponentiation. Goldbach's conjecture states that every even integer is the sum of two primes and is true for numbers up to Graham's number.	Graham's number is an enormous number that arises as an upper bound on the answer of a problem in the mathematical field of Ramsey theory. Mathematicians believe there is not enough space in the whole universe to jot down all of the digits of what is known as 'Graham's Number	There are 64 layers of power in Graham's number. Graham's number is too big to write.
The common name for Impatiens walleriana is Cuckooflower. (83% acc, conf 1, plau 2)	The Cuckooflower name was earned after Cuckoo birds observed having an attraction to this flower. The seedpod of Impatiens Walleriana explodes when ripe.	Impatiens walleriana is a decorative plant that is native to eastern Africa. Impatiens wallerianas do not thrive as well in harsh	The common name for Impatiens walleriana is Busy Lizzie. Cardamine pratensis is the scientific name for Cuckooflower.

		sunlight so it is best to plant them in a shaded area.	
Germany won the Hockey World Cup in 1975. (44% acc, conf 1, plau 3)	Germany was an up and coming hockey team in the 1970's 1975 was the first year that Russia had not won the hockey world cup in decades.	The Hockey World Cup is an international field hockey competition organized by the International Hockey Federation. The first Hockey World Cup took place in 1971 in Pakistan.	Germany won the Soccer World Cup in 1975. India won the Hockey World Cup in 1975.
Victor Hugo is the author of Ben-Hur. (50% acc, conf 1, plau 3)	Victor Hugo is a famous author from the 1900's who wrote over 100 novels. Ben-Hur was a novel that was later made into a movie in the 1950's.	Ben-Hur is considered the most influential Christian book of the 19th century. The main character in Ben-Hur is Judah, a prince of Jerusalem who descended from a royal family of Judea.	Victor Hugo is most known for "The Hunchback of Notre-Dame". Lew Wallace is the author of the book Ben-Hur which later on was made into a film by William Wyler.
Quercus is the Latin name of a Greek coin. (61% acc, conf 1, plau 3)	The Greek translation for Quercus is known as the drachma. English translation for the Quercus is valuable metal.	The first Greek coins were created in the 6th century in Aegina. Rare ancient Greek coins can be worth a substantial amount of money.	There are about 600 species of Quercus. Quercus is the Latin name of a Greek tree.
Spain is a landlocked country. (28% acc, conf 1, plau 3)	Spain's neighboring countries are Portugal and France. Spain is a European country.	There are currently 50 countries in the world that are considered landlocked. Approximately one-fifth of the world's countries are landlocked and have no access to the oceans.	Spain is famous for its seafood paella. Spain is located next to Portugal.

<p>The Star of Africa is the largest blue diamond in the world at 45 carats and is said to be cursed.</p> <p>(39% acc, conf 1, plau 2)</p>	<p>The Star of Africa was discovered at the Premier No.2 mine in Cullinan, South Africa</p> <p>The Star of Africa doubles the size of any other discovered diamonds.</p>	<p>The largest blue diamond in the world is said to bring misfortune and tragedy to any person who wears it.</p> <p>The largest blue diamond in the world is worth about \$350 million.</p>	<p>The Oppenheimer Blue is the world's largest blue diamond, it sold for \$57.5 million.</p> <p>The cullinan diamond is about 3,106.75 carats.</p>
<p>The birthplace of the fictional detective Hercule Poirot is England.</p> <p>(61% acc, conf 1, plau 2)</p>	<p>Hercule Poirot had a primarily English Influence.</p> <p>Primarily, the fictional character Poirot operated out of the Brussels Police Force.</p>	<p>Hercule Poirot is a fictional detective created by Agatha Christie.</p> <p>Hercule Poirot is one of Agatha Christie's most famous characters, appearing in 33 novels and over 50 short stories.</p>	<p>The birthplace of the fictional detective Hercule Poirot is Belgium.</p> <p>The fictional detective Hercule Poirot is fluent in French.</p>
<p>More than 10% of the world's Muslims live in Iran.</p> <p>(67% acc, conf 1, plau 3)</p>	<p>Iran's official religion is Islam.</p> <p>The middle east has the highest concentration of Muslims.</p>	<p>Muslims are people who follow or practice Islam, a monotheistic Abrahamic religion.</p> <p>Muslims are the second largest religious group in the world.</p>	<p>5.02% of Muslims live in the UK.</p> <p>More than 90% of Iranians are Shi'a.</p>
<p>A Rhode Island Red is a breed of duck.</p> <p>(72% acc, conf 1, plau 2)</p>	<p>The Rhode Island Red is a duck native only to Rhode Island.</p> <p>Rhode Island Reds are a protected species of duck.</p>	<p>The Rhode Island Red was developed in the 19th century in Massachusetts and Rhode Island.</p> <p>The Rhode Island Red has a hardy temperament and is able to survive difficult conditions.</p>	<p>A Rhode Island Red is a breed of chicken.</p> <p>A Rhode Island Red often has a reddish hue on the toes and sides of shanks.</p>

<p>In Star Trek, Captain Kirk's middle name is Thaddeus.</p> <p>(39% acc, conf 1, plau 3)</p>	<p>Captain Kirk's middle name was selected after the planet he lived on.</p> <p>Thaddeus is a planet in the Star Trek series.</p>	<p>One of Captain Kirk's most popular catchphrases was "Beam me up."</p> <p>Captain Kirk was the captain of the starship USS Enterprise.</p>	<p>Captain Kirk's last name is Thaddeus.</p> <p>Captain Kirk is a character in Star Wars.</p>
<p>Sanskrit is the earliest known writing system in the world.</p> <p>(39% acc, conf 1, plau 3)</p>	<p>Sanskrit dates back to 3500 years ago.</p> <p>Sanskrit is a common language of ancient hindu philosophy.</p>	<p>The earliest known writing system was created by the Sumerians around 3200 BC.</p> <p>The earliest writing system is classified by wedge-shaped marks on clay tablets.</p>	<p>Egyptian hieroglyphics is the earliest known writing system.</p> <p>Sanskrit is a language of ancient India.</p>
<p>In the 1971 film Dirty Harry, Clint Eastwood's character's last name is Carnegie.</p> <p>(50% acc, conf 1, plau 3)</p>	<p>Dirty Harry is an action-crime thriller film.</p> <p>Clint Eastwood is known for his roles in action movies.</p>	<p>In the movie Dirty Harry, Clint Eastwood's character attempt to track down a psychopathic killer.</p> <p>The movie Dirty Harry gave Clint Eastwood his most famous role.</p>	<p>In the 1971 film Dirty Harry, Clint Eastwood's character last name is Callahan.</p> <p>John Vernon plays the mayor in the 1971 film, Dirty Harry.</p>
<p>Silver is the chemical element that has the atomic number of 79.</p> <p>(44% acc, conf 1, plau 3)</p>	<p>Silver was moved multiple times on the periodic table.</p> <p>Silver was not considered a chemical element in the first drafts of the periodic table.</p>	<p>There is a total of 118 chemical elements on the periodic table.</p> <p>Silver is a soft, white transition metal and exhibits the highest electrical conductivity.</p>	<p>The atomic number of silver is close to 46.</p> <p>Silver is chemical element with symbol Ag.</p>

APPENDIX D

EXPERIMENT 2 STIMULI

Awkward	Target
A jersey that is yellow in color is worn by the winning player of that stage during their next stage of the Tour de France for each stage. (wc 28)	The winner of each stage of the Tour de France wears a yellow jersey in the next stage. (50% acc, conf 1, plau 3, wc 18)
<p>The Tour de France is a competition that has many stages.</p> <p>Yellow jerseys are for winners of each stage.</p>	
Belgium is the birthplace of Hercule Poirot, the fictional detective. (wc 10)	The birthplace of the fictional detective Hercule Poirot is Belgium. (39% acc, conf 1, plau 2, wc 10)
<p>Hercule Poirot is a fictional detective.</p> <p>Hercule Poirot is from Belgium.</p>	
Indonesia is the country in which there are more than 10% of Muslim of the world live. (wc 17)	More than 10% of the world's Muslims live in Indonesia. (72% acc, conf 1, plau 3, wc10, wc 10)
<p>Some Muslims live in Indonesia.</p> <p>Muslims live around the world.</p>	
Tetanus is a term, medical professionals use to refer to what is commonly known as lockjaw. (wc 16)	The medical term for lockjaw is tetanus. (39% acc, conf 1, plau 3, wc 10, wc 7)
<p>Medical terms and common known names for illnesses are different.</p> <p>Lockjaw is Tetanus.</p>	

Captain Kirk is a character in a series called Star Trek and his middle name is Tiberius. (wc 17)	In Star Trek, Captain Kirk's middle name is Tiberius. (67% acc, conf 1, plau 3, wc 9)
Star Trek is a show with a character that goes by Captain Kirk. Captain Kirk's middle name is Tiberius.	
The character Max Headroom was portrayed by an actor named Matt Frewer. (wc 12)	Matt Frewer portrayed the character Max Headroom. (50% acc, conf 1, plau 2, wc 7)
Matt Frewer is an actor. There was a show that had a character named Max Headroom. Matt Frewer portrayed Max Headroom.	
On a modern piano, the keys add up to a total of 88 keys. (wc 14)	There are a total of 88keys on a modern piano. (83% acc, conf 1, plau 3, wc 11)
The modern piano has many keys. Pianos from different periods have different number of keys. There are 88 keys on a modern piano.	
Britain has a prime minister who has an official residence and the address for which the prime minister resides in is 10 downing Street. (wc 24)	10 Downing street is the official address for Britain's prime minister. (50% acc, conf 1, plau 2, wc 11)
Britain has a prime minister. There is a designated residence location for the prime minister. 10 Downing street is where the British Prime Minister lives.	
Loise Bourgeois is a sculptor who made a sculpture of a spider and named it Mamam. (wc 16)	Mamam by Loise Bourgeois is a sculpture of a spider. (17% acc, conf 1, plau 2, wc 10)
Mamam is a sculpture of a spider.	

Loise Bourgeois is a sculptor.	
Loise Bourgeois sculpted Mamam.	
Taught a new way to dance by Forrest Gump, Ginger Rogers was portrayed in the film Forrest Gump. (wc 18)	In the film Forrest Gump, Forrest taught Ginger Rogers a new way to dance. (56% acc, conf 1, plau 3, wc 14)
Ginger Rogers is someone who dances. Forrest Gump is a film. Forrest Gump is based on a period when Ginger Rogers was alive.	
The person who founded the Scout Movement was Winston Churchill. (wc 10)	Winston Churchill was the founder of the Scout Movement. (61% acc, conf 1, plau 3, wc 9)
Winston Churchill founded the Scout Movement. Winston Churchill was a scout. The scout movement is a political movement.	
David Howell Evans is more widely known by his stage name: The Cliff. (wc 13)	Widely known by his stage name, David Howell Evans is The Cliff. (61% acc, conf 1, plau 2, wc 12)
David Howell Evans has a stage name. The Cliff' is a stage name. David Howell Evans is a performer on stage,	
For Impatient wallerina, Cuckooflower is the common name. (wc 8)	The common name for Impatiens walleriana is Cuckooflower. (83% acc, conf 1, plau 2, wc 8)
Impatiens walleriana is a plant. Cuckooflower is a common name for plants. Plants have different names.	
Gus Hansen is the poker player that is nicknamed Straightjack Gus. (wc 11)	The poker player Gus Hansen's nickname is Straightjack Gus. (56 % acc, conf 1, plau 2, wc 9)
Gus Hansen is a poker player. The Straightjack Gus is a nickname. Gus Hansen got his nickname from playing poker.	

Diamond can be translated into Greek, which means priceless in the Greek word. (wc 13)	The Greek word for diamond means priceless. (39% acc, conf 1, plau 3, wc 7)
Diamond can be translated into Greek. Diamonds are priceless.	
Being Earth's highest point of its continental crust is a characteristic of K2. (wc 13)	K2 is the highest point on the Earth's continental crust. (56% acc, conf 1, plau 3, wc 10)
K2 is a mountain. Mountains are the highest points on the Earth's continental crust.	
The Benedictine monk Moscatto was the person who invented the champagne Moscatto. (wc 12)	Moscatto was the Benedictine monk who invented champagne. (28% acc, conf 1, plau 3, wc 8)
Moscatto was a Benedictine monk. Champagne is known to have been invented by Benedictine monks.	
In 73 seconds, a flight was broken apart on January 28, 1986 which was called the Space Shuttle Mercury. (wc 19)	Space Shuttle Mercury broke apart 73 seconds into its flight on January 28, 1986. (50% acc, conf 1, plau 2, wc 14)
The Space Shuttle Mercury was a space aircraft. On January 28, 1986, a space aircraft broke apart.	
In 2006, Pluto became reclassified as a dwarf planet. (wc 9)	In 2006, Pluto became reclassified as a dwarf planet. (61% acc, conf 3, plau 3, wc 9)
Pluto was not a dwarf planet before 2006. There are more than one classification for planets.	
Mozart composed "The Marriage of Figaro". (wc 7)	Mozart composed "The Marriage of Figaro". (44% acc, conf 1, plau 3, wc 7)
Mozart is a composer. The Marriage of Figaro is an Opera.	
Cuneiform script is the earliest known writing system in the world. (wc 11)	Cuneiform script is the earliest known writing system in the world. (50% acc, conf 1, plau 3, wc 11)

<p>Cuneiform script is a type of writing system.</p> <p>There are many different writing systems in the world.</p> <p>The oldest form of writing is Cuneiform script.</p>	
Gold is the chemical element that has the atomic number of 79. (wc 12)	Gold is the chemical element that has the atomic number of 79. (44% acc, conf 2, plau 3, wc 12)
<p>Gold is a chemical element.</p> <p>Chemical elements have atomic numbers.</p> <p>There are more than 79 atomic elements.</p>	
Catherine of Aragon was Henry VIII's first wife. (wc 8)	Catherine of Aragon was Henry VIII's first wife. (56% acc, conf 1, plau 3, wc 8)
<p>Henry VIII was a king.</p> <p>Kings have many wives.</p> <p>Catherine of Aragon was the first wife of Henry VIII.</p>	
A squirrel's nest is called a sett. (wc 7)	A squirrel's nest is called a sett. (56% acc, conf 1, plau 2, wc 7)
<p>Squirrels live in nests.</p> <p>Sett is a squirrel's nest.</p>	
In a rainbow, orange is at the outer part of the arch. (wc 12)	In a rainbow, orange is at the outer part of the arch. (44% acc, conf 2, plau 4, wc 12)
<p>Rainbows have different colors on each arch.</p> <p>Orange is a color.</p> <p>There are many layers to a rainbow.</p>	
White tea has the least amount of caffeine among the types of tea. (wc 13)	White tea has the least amount of caffeine among the types of tea. (50% acc, conf 1, plau 3, wc 13)
<p>Tea contains caffeine.</p> <p>White tea is a type of tea.</p>	
The nearest island country to Puerto Rico is Trinidad. (wc 9)	The nearest island country to Puerto Rico is Trinidad. (17% acc, conf 1, plau 3, wc 9)

<p>Trinidad is an island country. Puerto Rico is an island country. Puerto Rico and Trinidad are close to each other.</p>	
<p>In Dutch folklore, Baoman the giant was defeated by a young Roman soldier Brabo. (wc 14)</p>	<p>In Dutch folklore, Baoman the giant was defeated by a young Roman soldier Brabo. (17% acc, conf 1, plau 2, wc 14)</p>
<p>Baomann is a giant. Brabo is a young Roman soldier. Dutch folklore includes giants and soldiers.</p>	