THE INTERACTIONS OF GENERATION AND SELF-REFERENCE

INTERACTIONS BETWEEN THE GENERATION AND SELF-REFERENCE EFFECTS

By YUTIAN MA, B.SC. (HONS.)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements

for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2025) Hamilton, Ontario (Psychology, Neuroscience & Behaviour)

TITLE: Interactions between Generation and Self-Reference Effects

AUTHOR: Yutian Ma, B.Sc. (Hons) (McMaster University)

SUPERVISORS: Dr. Joseph A. Kim & Dr. Faria Sana

NUMBER OF PAGES: xii, 55

Lay Abstract

This study investigated two well established strategies for improving memory: the generation effect, where information is better remembered when individuals produce it themselves, and the self-reference effect, which enhances memory by relating information to oneself. We asked participants to complete sentence fragments like “Your glass contained your ___” or “Their glass contained their ___” using either self-related or other-related wording. We also compared how well they remembered the sentences using two types of memory tests: free recall (remembering as much as you can) and cued recall (remembering with hints).

We found that generating answers significantly improved memory. We also found that using “your” instead of “their” also gave a small boost to memory, showing that even subtle language cues can make information feel more personal and therefore more memorable. Interestingly, these two strategies didn’t interfere with each other, meaning people can use both at the same time to provide unique benefits to learning. These findings have practical implications for education and communication, demonstrating that both active engagement and subtle shifts in language can meaningfully enhance memory.

Abstract

This thesis explored the individual and combined effects of two well-established memory-enhancing strategies: the generation effect and the self-reference effect. A total of 89 undergraduate participants completed sentence completion tasks that manipulated both generative tasks (generate vs. read) and personal relevance through the use of second-person possessive pronouns (“your” vs. “their”). Memory performance was assessed using both free and cued recall tasks to evaluate how these encoding strategies operate across different forms of relational memory. A 2x2 mixed ANOVA revealed a main effect of generation, with the generation condition recalling significantly more items than the read condition across both memory tests. A smaller but statistically significant main effect of self-reference was also observed, with items from sentences with self pronouns (“your”) being recalled significantly more than items from sentences with other pronouns (“their”). This suggests that even minimal linguistic cues can elicit enhanced encoding when linked to the self. Importantly, the interaction between generation and self-reference was not significant, indicating that these strategies provide additive benefits and can be used concurrently without interference. An exploratory analysis revealed that the generation effect was significantly larger in cued recall compared to free recall, possibly due to stronger cue-target associative encoding. These results contribute to our understanding of the underlying mechanisms of memory and support the practical application of combining generative learning with personalized language. The findings have implications for educational and cognitive training contexts, where both strategies may be integrated to enhance memory performance in a variety of learners.

Acknowledgments

I would like to first extend the warmest and most sincerest of gratitudes to my supervisors, Dr. Joseph Kim and Dr. Faria Sana for all their support, advice, and encouragement. Thank you for providing me with the opportunity to conduct research on a topic that I was truly passionate about. Over these past few years, I hope that I have shown you how I have grown both as an academic and as a person due to your influences. I would also like to thank the other members of my supervisor committee, Dr. Bruce Milliken and Dr. Daniel Goldreich. You have both made many insightful comments and suggestions that greatly impacted my final thesis. I would also like to thank Dr. David Shore for suggesting a yoked design, which greatly reduced the headache I was having at the time. Finally, I wish to extend my gratitude to Dr. Lina Deker for her help in conceptualizing and designing the study. Without you, I would have made so many mistakes and I am so glad to have had your support. To all my current and previous lab mates, I would like to thank you all for all the fond memories and productive lab meetings. Thank you for always listening to my comments and opinions even when they came from an old man. A special thanks to Veronica Cui for always providing feedback and making sure that I meet deadlines. I have learned a lot from you, especially about “girly pop” and “Taylor Swift”. To the students who worked with me directly, Adjoda Archer, Alishba Ahmed, and Jessica Hur, thank you for all the time and effort you put into this project. I greatly appreciated all the coding, writing, and participant running that you all did. Finally, to all my friends, family, and loved ones, I would like to express my absolute highest level of thanks for all the memories we made together. Even if you weren’t cognizant of it, your presence made my life better and greatly assisted me while I was writing this thesis. Once again, I would like to thank you all as I hope to honor your unwavering support by completing this thesis.

Table of Contents

Introduction 1
Literature Review 2
  Generation 2
  Self-Reference Effect 5
  Interactions 8
The Present Study 12
  Study 12
  Implications 13
  Methods 14
    Participants 15
    Materials 16
    Creating the Yoked Design in the Read Condition 18
    Counterbalance 19
    Learning phase 19
    Distractor task 20
    Measures 20
    Scoring 20
Results 21
    Demographics 21
    Statistical Analysis 21
    Outliers 22
    Assumption Checks 22
    Data Visualization and Confidence Intervals 23
    Cued Recall 23
      Post Hoc Comparisons 24
  Free Recall 26
      Post Hoc Comparisons 26
  Summary 28
  Recall Condition 28
Discussion 30

Generation Effect 30
Interactions 31
Self-Referential Possessive Pronouns 33
Practical Implications 36
Limitations and Future Directions 37
Conclusions 39
References 41
Appendix 50

List of Figures

Figure 1A: Bar Graph of Cued Recall Scores Across Generate and Self Conditions 26
Figure 1B: Bar Graph of Free Recall Scores Across Generate and Self Conditions.....................28
Figure 2: Line Graph Showing the Generate Condition–Recall Test Interaction 30
Figure 3A: Q–Q Plots Assessing the Normality of Residuals within the Cued Recall Scores Across the Generation and Self Conditions...53
Figure 3B: Q–Q Plots Assessing the Normality of Residuals within the Free Recall Scores Across the Generation and Self Conditions...55

List of Tables

Table 1: Table of Subject Nouns and Verbs Used to Create Sentences Along with their Difficulty to Understand Scores and Valence Scores..50

List of Abbreviations and Symbols

ICC, *interclass correlation*

ges, *generalized eta squared*

ANOVA, *analysis of variance*

SD, *standard deviation*

IQR, *interquartile range*

Q–Q, *quantile–quantile*

Declaration of Academic Achievement

I, Max Yutian Ma, declare that I am the sole author of this thesis. I was responsible for conceptualizing the research questions and conducting data analyses.

Dr. Joseph A. Kim, Dr. Faria Sana, Dr. Daniel Goldreich, and Dr. Bruce Milliken contributed to the design of this study by providing valuable suggestions and feedback throughout the research process.

Dr. Lina Deker provided valuable input in the conceptualization of the study and designing of both the methods and the final sentences that we used. In addition to their help running the study, Adjoda Archer, Alishba Ahmed, and Jessica Hur all contributed to narrowing down the final sentences as well as how to grade the recall scores.

Introduction

Over the past few decades, extensive research has been conducted on the underlying mechanisms of forming memories. Through encoding, storage, and retrieval, there have been many developments made towards understanding the intricacies of each step of the memory process (Guskjolen, 2023; Tulving & Thomson, 1973). Various factors have been found to affect subsequent memory quality. These factors can broadly be classified into two groups: item specific factors, which are related to distinct details of the item being encoded, and relation memory factors, which refer to indirect associations between items. Item specific factors such as depth of processing enhance the encoding phase by processing deeper semantic information, resulting in stronger memory formation (Craik & Tulving, 1975). Additionally, attentional resources during encoding and distinctiveness also play a part in strengthening item specific memory (Cowan et al., 2024; Hunt, 2003). On the other hand, relation memory factors like associative strengthening describe how contextual cues during the encoding phase can then become aids in the retrieval phase of memory, with richer and more congruent connections resulting in better memory (McCurdy et al., 2020). All of these factors contribute to the multifactor theory of memory, which describes how memory performance is the result of numerous and distinct factors (Hirshman & Bjork, 1988).

As research on these factors naturally progressed, an increasing number of mnemonic techniques were also defined. Accordingly, researchers then refined these techniques into distinct strategies, applicable in various real world scenarios. Instructors and students in particular, serve to benefit greatly from understanding and employing strategies that can help learners study and memorize information more efficiently. Among these strategies, the generation and self-reference effects are particularly beneficial to students. Understanding both are crucial in our current understanding of memory and how we continue to study memory.

Literature Review

Generation

The generation effect describes an improvement to the memory of self-generated information by the learner during the encoding process. This is in contrast to information that is read or passively encoded. As generation is a broad concept, numerous activities are considered when discussing generative strategies. Of note, the pioneering study by Slamecka and Graf (1978) explored the benefits of generation through word pairs. Participants had better memory of target words if they generated the word (e.g., Water-W___) over reading the word (e.g., Water-Wet). These effects extend to sentences, where memory of the last word of a sentence was better when the participant generated it (e.g., The dove is a symbol of _____) compared to reading the completed sentence (Kane & Anderson, 1978). Ghatala (1981) expanded on this sentence completion task by adding a judgment condition in which participants read a completed sentence and determined if the last word was semantically appropriate. While both conditions performed better than the read condition, the semantic judgement condition did not significantly differ from the generated condition. This suggests the benefits of the generation effect may be due to a deeper semantic level of encoding. However, another study done independently by McFarland et al. (1980) used similar methods and found the generate condition to significantly improve memory over the judgement condition. This discrepancy may be explained by the nature of the sentences they used to prompt generation. In the McFarland et al. study, the target word was not constrained by the experimenter, meaning there were multiple viable words that could be used to complete the sentence (e.g., "Explorers discovered the ______ fifty years before the settlers came."). In contrast, Ghatala's study sentences were designed to constrain the generated word to just one. These two studies in conjunction suggest an effect of constraint during generation, in which participants gain a greater benefit to memory when they have more freedom in what they generate.

A more recent study by McCurdy et al. (2017) explored this effect of constraint on the generation effect. They used word pairs to compare low and high constraint generation. Ultimately, lower constraint generation, where participants freely generated any target word they wished, led to better memory when compared to higher constraint generation, where participants would be forced to generate a specific target word. This finding was further expanded in a meta-analysis in which experiments that strictly controlled generation showed overall lower magnitudes of the generation effect when compared to experiments that gave little or no directions on how to generate (McCurdy et al., 2020).

The effects of constraint extend to the application of generative activities, in which large effect sizes are found in free generation activities. One such activity is note-taking, where students who wrote notes on a passage better remembered its contents than those who simply read it (Doctorow et al., 1978). Consistent with the effects of constraint on generation, there is a stronger benefit for memory when participants are allowed to paraphrase or summarize the contents in their own words over direct copying (Bretzing & Kulhavy, 1979; Coleman et al., 1997). It has also been shown that generating test questions also served to improve comprehension test scores, although these activities only benefited learners when generating and answering conceptual questions, not direct retrieval questions (Bugg & McDaniel, 2012). This is consistent with the effects of constraint seen thus far as participants have more options in how they approach a conceptual question but are limited in how they word retrieval questions.

Low constraint may improve memory through imagery, which describes how some generative activities encourage the generation of internal visual representations. An obtuse example is mapping, in which learners create a spatial arrangement of related ideas or words. Usually in the form of a diagram or graphic, these maps represent the structure of connections between ideas in the form of causes, effects, hierarchies and comparisons. Students who study concepts by creating such mind maps perform better on memory tasks than students who just take direct notes (Holley et al., 1979; Ponce & Mayer, 2014). Another generative activity which improves memory through imagery is drawing an illustration depicting the contents of a lesson (Leutner et al., 2009; Schwamborn et al., 2010). Against the prior discussion on constraint, some guidance on how to draw is needed or else memory performance decreases. However, this might be due to the learner using some cognitive load on the mechanisms of drawing, rather than on encoding the contents of the lesson (Leutner & Schmeck, 2014). Leopold and Mayer (2015) addressed some of these concerns by removing the physical aspect of drawing and instead relying solely on imagining as a form of generation. They found that participants who were instructed to generate a mental image of the human respiratory system did better on memory tests than those who just read the passage. Furthermore, the effect diminished when participants were given a picture of what to imagine. While the authors suggested this may be due to a lesser incentive to produce a mental image, it may have also been due to the picture constraining the participant's generation of the mental image.

The generation effect has been repeatedly shown to become more potent in lower constraint conditions. While some theories have been raised as to why, we posit that an explanation may be found by considering the self-reference effect.

Self-Reference Effect

The self-reference effect describes an increase in memory specifically for information highly related to the learner. In classic studies of the self-reference effect, participants would judge an item, usually a word or sentence, on a scale of relatedness to themselves (e.g., Does this word apply to you?). Then, researchers would find that items judged high in relatedness would be better remembered on subsequent memory tests compared to stimuli judged to be of lower relatedness (Cunningham et al., 2014; Klein & Kihlstrom, 1986; Rogers et al., 1977; Symons & Johnson, 1997). This increase in memory performance is not just due to semantic processing. When participants were faced with both a relatedness judgement condition (e.g., Does the word "cat" apply to you?) and a semantic judgement condition (e.g., Does the word "cat" relate to "dog"?), the self-related items were better remembered than the semantic judgement items. (Leshikar et al., 2015). This was especially true for items that were seen as positive traits, which is further evidence that the self-reference effect influences processing beyond just deeper semantics.

Emotional valence is a modulator of the self-reference effect. Valence describes the general emotional affect of an item and can be positive, neutral, or negative. For example, "the puppy wagged its tail" would be considered a positive valence sentence as it is likely to elicit positive emotions. On the contrary, "the puppy begged for food" is a negative valence sentence as it is likely to elicit more negative emotions (Sedikides & Green, 2004). Valence is important to consider in self-referential studies because participants are more likely to remember self-related positive valence descriptors over negative valence descriptors (D'Argembeau et al., 2005). The effects of valence do not apply to descriptors judged to be other-related. This is thought to result from the cognitive biases inherent in self-processing. Individuals tend to hold positively biased self-schemas, which promote deeper encoding and greater elaboration of self-relevant positive information compared to negative information (Sedikides & Green, 2000). This self-enhancement motive facilitates memory advantages for positive content that aligns with one's self-concept, whereas negative information may be neglected during encoding. In contrast, when processing information about others, especially unfamiliar others, such motivational biases are weaker or absent, leading to more balanced encoding of positive and negative descriptors (Fields & Kuperberg, 2012). As a result, positive valence self related items tend to be processed more deeply, amplifying the self-reference effect through depth of processing factors (D'Argembeau et al., 2005).

While the self-reference effect is typically strongest when information is processed in direct relation to the self, emerging evidence suggests that similar memory benefits can occur for information on individuals related to participant. Family members or close friends may be partially integrated into one's self-concept, allowing them to also benefit from self-referential processing (Aron et al., 1991). Another study by Bower and Gilligan (1979) found that referencing close friends during encoding improved recall compared to semantic processing, though not to the same degree as self-referencing. Generally, reporting actions that reference a familiar other (e.g., Recall an incident in which YOUR MOTHER exemplified this trait) has comparable memory benefits to referencing oneself (Recall an incident in which YOU exemplified this trait; Klein et al. 1989). This suggests that the self-reference effect is robust and can be induced through familiarity. Of note, this familiarity tends to be denoted through second person pronouns.

Beyond familiar others, possession has also been shown to enhance memory for objects. In an experiment by Sui and Humphreys (2013), a participant was shown a series of items that were arbitrarily assigned to belong to either the experimenter or themselves (This is YOUR ball vs This is MY ball). Participants then had better memory of the properties of the items "belonging" to them. In another experiment, they instructed participants to associate shapes with a personal label (e.g., you, friend, stranger). Afterwards, they presented participants with shape-person pairs and had participants make judgments on whether these pairs matched their earlier associations. Shapes that were associated with the self (you) were judged faster than shapes associated with others (friend, stranger) (Sui et al., 2012).

This advantage in processing personal pronouns is maintained in other processing activities. Participants are faster and more accurate when solving application-based math problems written with the self-pronoun "you" compared to when the problems used the names of other people (Cunningham et al., 2024). This is due to how the self plays a central part of human processing, functioning as the most accessible point of reference in new experiences (Banaji & Prentice, 1994; Hood, 2012). As memory encoding relies on perception, it follows that the self is integral to how we perceive information and turn that information into memories (Sui & Humphreys, 2015). Thus, we reason that the framing of information in reference to either self or other would yield different memory outcomes. An experiment by Ditman et al. (2010) had participants read sentences containing actions (e.g., I am slicing the tomato) or descriptions (e.g., I am a 36 year old barber). These sentences used either "you", "I' or "he" as pronouns. Afterwards, memory tests showed that participants had better memory of sentences with the self-pronoun ("you") compared to the other-pronouns ("I" and "he") but only for sentences describing actions. As for the rationale, there is evidence that participants better remember these "you" pronoun sentences because we process the world in reference to the self. Participants likely form a mental image during the encoding phase in which the richness of this image is directly associated with memory performance (Nilsson, 2000; Pavio, 1971; Zimmer et al., 2001). Thus, the "you" pronoun sentences produce the most salient mental image due to participants imagining a sentence with themselves as the subject (Brunyé et al., 2009). This would also explain the discrepancy between action statements and descriptive statements as some descriptions would not accurately describe the participants in self-pronoun sentences, making the mental image incongruent with reality. Additionally, addressing the learner through declarative statements which are naturally in second person also increases their learner engagement, which likely also contributes to a memory benefit (Qui & Jiang, 2021).

The self-reference effect is a robust and well-studied phenomenon that results in deeper level processing. This is one of many comparisons that one could make between the self-reference effect and the generation effect.

Interactions

The present study aims to explore whether there is an interaction between the generation and self-reference effects. Specifically, we want to determine if the benefits in memory follow either (1) an additive model, in which each factor contributes a fixed gain regardless of the other, or (2) an interactive model, in which the effects self-reference disproportionally boost memory in the read condition over generate. We reason that there is likely to be some interaction as self-generated information is more likely to be related to the learner, thus exhibiting both generation and self reference effects. This would be further applicable to low-constraint generation conditions as the generated material would be more likely to be personalized to the participant when they are free to generate whatever they want.

Additionally, there is some overlap in the theories of the underlying mechanisms between the two effects. One theory is depth of processing, in which information that is more semantically relevant is better remembered (Craik & Tulving, 1975). Both the generation and self-reference effects can influence deeper semantic levels of processing (Klein & Loftus, 1998; McCurdy et al., 2020; Symons & Johnson, 1997). Likewise, both effects benefit from emotional valence, although how they interact with valence differs. As stated, the self-reference effect has a positive bias in which positive valence receives a disproportionately larger benefit to memory (D'Argembeau et al., 2005; Sedikides & Green, 2004). On the other hand, the interactions between valence and the generation effect have been mixed. While words with higher levels of emotionality enhanced the generation effect, there was no difference between positive and negative valanced words (Berrin-Wasserman et al., 2003). This manifests as a benefit to item-specific memory, in which there is an enhanced ability to discriminate the target item from other similar distractors (Hunt & Einstein, 1981). Additionally, both effects may enhance memory by improving the ability to mentally visualize the content being encoded. Outside of imagining as a generative activity, generating information requires active construction, which can lead to richer mental imagery (Leopold & Mayer, 2015). Similarly, self-referential encoding often engages personally relevant schemas, making it easier for participants to imagine themselves in the described scenarios (Brunyé et al., 2009). This may suggest that the two effects both tap into mechanisms of imagery when providing their memory benefits.

In a similar vein, both effects show enhancements for relational memory, in which the contextual cues strengthen the ability to properly recall (He et al., 2021; Hirshman & Bjork, 1988). As such, one theory for the generation effect is that it strengthens contextual memory, in that the context during encoding (physical qualities of text, who performed the action) is more salient and could potentially assist memory by acting as a cue during recall (Geghman & Multhaup, 2004; Marsh, 2006). Although these findings have been mixed, it is generally agreed that generation improves conceptual details like the source of the memory over physical details like font size and location (Mulligan, 2011). Likewise, the self-reference effect has been shown to improve relational memory as well. In particular, source memory and metamemory judgements of source memory both show enhancements. When asked to either generate or watch another person generating a word, participants were able to correctly predict that they would remember more self-generated items over other-generated items (Carroll et al., 2001). When considered together, these studies suggest that the effects of generation and self-reference may use similar memory processes and therefore may interact when presented in tandem.

This culminates into the multifactor theory of memory, which posits that successful memory is dependent on multiple distinct factors (Hirshman & Bjork, 1988; McCurdy & Leshikar, 2022). One method of parsing out the specific factors involved is through investigating a multitude of testing conditions. For example, item-specific memory is mainly tested during recognition tasks while relational memory is mainly tested in free and cued recall tests (Hunt & McDaniel, 1993). Some theories separate relational memory into cue-target and intertarget relational memory, with cue-target memory tested by cued recall and intertarget memory tested by free recall (McDaniel et al., 1990). The generation effect has been shown to improve performance on recognition, free recall, and cued recall tests, which supports this multifactor theory. One particular pattern to note is the larger effect sizes in cued recall tests over free recall, which suggests that intertarget memory does not receive as big a benefit as the other factors

(McCurdy et al. 2020). A meta-analysis of self-reference effects found similar results, in which all three types of tests benefit from self-reference over other-reference during encoding. In these methodologies where self and other conditions are compared, a similar pattern to the generation effect is found where the effect size tends to be smaller in free recall when compared to cued recall (Symons & Johnson, 1997). As these two effects show similar patterns within the factors of memory, we reason that this may be further evidence that they share some underlying mechanisms when conferring their benefits to memory.

In summary, the evidence suggests a potential interaction: as learners generate information, that material becomes increasingly self-relevant. This should be especially true under lower constraint generation conditions in which participants are free to generate whatever they want. Yet despite their conceptual overlap, empirical work has almost always examined the generation and self-reference effects in isolation. The present study has two main goals. (1) First, we wish to investigate whether the cognitive advantages of the generation and self-reference effects are orthogonal and therefore additive OR if there is an interaction. Clarifying this relationship is pivotal for refining models that attribute both effects to shared item-specific and relational processes and for designing learning activities that take advantage of the two effects. (2) We also want to investigate the viability of using second person in combination with possessive pronouns (e.g., your) to induce a self-reference effect, which would also have interesting real-world applications. The next section describes the experimental method we used to investigate these two goals.

The Present Study

Study

The present study seeks to examine the relationship between a generative activity (completing a sentence by generating the last word) and the influence of self-referential possessive pronouns (your vs their). As stated earlier, we sought to answer two main research questions.

1: To examine if there is an interaction between the generation and self-reference effects.

2: To investigate the effectiveness of self-reference on memory, by exploring memory differences in using second person possessive pronouns (your) versus third person possessive pronouns (their).

To address the first question, we used a mixed 2x2 design with the between subjects variable generation (generate vs read) and within subjects variable self (self vs other). The task was similar to the Ghatala (1981) in which participants would be tested on their recall memory for a target word within a cue sentence. This target word would be generated in the condition and read in the read condition. The cue sentences would contain either self-referential second person pronouns (your) in the self condition and other-referential third person pronouns in the other condition. This was done in order to detect any interaction effects between the two variables.

Based on previous research, we reason that there would be some overlap between the underlying memory processes driving the generation and self-reference effects. Thus, we hypothesize a significant interaction.

To address the second question, we examined the main effect of "self". Furthermore, we performed post-hoc pairwise t-tests between the self and other conditions in order to determine if there was a significant difference between groups. As there are many studies that have reported effects of possession and self on memory, we hypothesize that there is likely to be a significant main effect of the self variable.

Implications

The results of this study have important implications for both theoretical memory research and practical educational strategies. At a theoretical level, examining the interaction between the generation and self-reference effects offers valuable insight into how distinct cognitive mechanisms may overlap during the various phases of memory. If these effects are found to interact, it would suggest a shared or synergistic mechanism in which generative activities are naturally self-referential in nature. In addition, the presence (or absence) of an interaction would provide further insight into the cognitive processes that underlies both strategies and further refine the current understanding of the multi-factory theory of memory. If there is an interaction, then it lends support to the idea that item and relational factors may be enhanced through different techniques. If no interaction is found, then the results support a further division between the factors, showing how these memory mechanisms function separately.

Furthermore, the study contributes to a growing body of research that emphasizes the role of personalization in memory formation. If self-referential cues, such as second-person pronouns used in possessive framing, are found to enhance generative learning, this would provide further evidence that memory encoding is significantly influenced by the degree to which material is perceived as personally relevant. This has practical applications in domains such as education and instructional design, where learning outcomes could be improved by integrating self-relevant framing into generative tasks (e.g., customized learning prompts, personalized practice exercises, and learner-centered note-taking strategies).

Additionally, the exploration of second-person pronouns ("your") as a trigger for self-referential encoding provides a simple yet potentially powerful tool for improving memory in everyday contexts. If effective, this technique could be easily implemented in educational materials and instructional language without the need for complex interventions or assessments of personal relevance.

More broadly, this research highlights the importance of designing learning experiences that are both active and personally meaningful. By studying how these two effects interact to shape memory, the findings from this study contribute to the vast literature that may encourage educators and researchers to shift learning from passive content consumption to active, individualized engagement.

Methods

Our experiment was conducted using the survey software, Qualtrics. It consisted of a 2x2 mixed design with the between subjects variable being generation (generate vs read) and the within subjects variable being self (self vs other). In the generate condition, participants were presented with 48 sentences during the learning phase. Participants saw one sentence at a time for 15 seconds. These sentences were missing the last word, which served as the target word to-be-remembered (e.g., Your mother made your _______). Participants were asked to complete the sentence by generating whatever word they thought was appropriate. Half the sentences belonged to the self condition, in which the possessive pronouns were always "your" (Your mother made your _____) and the other half of the sentences belonged to the other condition, in which the possessive pronouns were always "their" (Their mother made their _____). In the read condition, participants instead read an already completed sentence and were instructed to type out the underlined last word, which served as the target to-be-remembered word (Your mother made your <u>bed</u>). To complete the sentences for the read condition, we used a yoked design in which we used a generated target word from each participant in the generate condition Afterwards, participants completed a distractor task before moving on to the testing phase. During the testing phase, they first completed a free recall test in which they tried to remember as many target words as possible. Afterwards, they were given the noun stems (e.g., Your mother) of each sentence and instructed to perform a cued recall test in which they tried to remember the target word they generated or read for each noun stem cue.

Participants

Undergraduate students enrolled in an Introductory Psychology course at McMaster University were recruited online through the McMaster SONA system. They were pre-screened to be English first language and have normal or corrected-to-normal vision and hearing. Participants randomly signed up for a time slot in our study and were instructed to attend an in-person study session. The final sample consisted of 89 participants (45 generate, 44 read). All procedures were approved by the McMaster Research Ethics Board (Protocol ID: 7246), and all participants provided informed consent before beginning the study and after the debriefing.

Due to the lack of previous literature describing an interaction between generation and self-reference using our specific methodology, we decided to run a simulation-based power analysis using the Superpower package in R (Lakens & Caldwell, 2021) to determine the appropriate sample size for detecting an interaction between the self-reference and generation effects. The study design was a 2 × 2 mixed ANOVA, with self condition (self vs. other) as a within-subjects factor and generation (generate vs. read) as a between-subjects factor.

Based on the theoretical predictions outlined above, we modeled a non-additive interaction where combining self-reference and generation would not yield a fully additive memory benefit. While there was extensive literature on the effect sizes of the generation and self-reference effects, there was no past literature on an interaction. Thus, we chose our means and standard deviations to aim specifically for a low interaction effect size (Cohen's F>0.15). The following cell means were specified to reflect this pattern: read–other: 0, read–self: 4, generate–other: 2, generate–self: 5. We also choose a standard deviation of 0.5 and a within-subjects correlation of 0.1.

With 40 participants per between-subjects group (N = 80 total), these values resulted in a large main effect of generate (power >99%, Cohen's F = 0.85) and a smaller main effect of self (power> 99%. Cohen's F = 0.560). More importantly, the simulation estimated a statistical power of >90% for an interaction with Cohen's F of 0.131, which was what we were specifically aiming for.

Overall, these results support that the planned sample size provides sufficient power (>80%) to detect the hypothesized interaction, while also producing well-powered main effects in line with previous literature.

Materials

Sentences were modeled after the sentences used in the Ghatala (1931) study. Over 140 sentences were originally created by the research team. Each sentence followed the specific formula (Possessive pronoun) + (Subject noun) + (Verb in past tense) + (Same possessive pronoun) + (Blank target noun). Care was taken to ensure that no subject noun or verb was repeated. To suit our research questions, we controlled for two variables. The first variable was emotional valence, which has been shown to consistently modulate the self-reference dimension (D'Argembeau et al., 2005; Sedikides & Green, 2004). We wrote half of the sentences with a positive valence (e.g., Your teacher enjoyed your _____.) and the other half with a negative valence (e.g., Your barista spilled your _____.). In order to ensure that the sentences were interpreted with the correct valence, seven independent blind raters were recruited to rate the positivity/negativity of the preliminary sentences. These same raters were also asked to rate how difficult it was to understand each sentence.

Valence was rated on a Likert scale from 0 to 10, with 0 representing "Not positive/negative at all", 5 representing "moderately positive/negative", and 10 representing "Extremely positive/negative". Difficulty was also rated on a Likert scale from 0 to 10, with 0 representing "not difficult to understand at all", 5 representing "moderately difficult to understand" and 10 representing "extremely difficult to understand".

Inter-rater reliability was assessed using a two-way random effects model (ICC(2,1); Shrout & Fleiss, 1979). The resulting ICC for valence was **0.35**, indicating moderate agreement among raters. Averaging the difficulty ratings across the seven raters yielded an estimated reliability of approximately **0.72**, based on the Spearman-Brown prophecy formula. The resulting ICC for difficulty was **0.17**, indicating low agreement among raters. Averaging the difficulty ratings across the seven raters yielded an estimated reliability of approximately **0.61**. For a full list of valence and difficulty ratings of each sentence, please refer to Table 1 in the Appendix.

The other variable we controlled for was the nature of the subject noun. Given that the self-reference effect exists for possession of both related persons and objects, we ensured that half of sentences contained object subject nouns (e.g., vase, phone, video etc.) and the other half contained person subject nouns (e.g., teacher, mother, baker).

Based on the valence and difficulty ratings, we narrowed down the sentences to 48. The final sentences were chosen to ensure a combination of moderate to high average valence rating (at least 4) and low difficulty rating (less than 3) while maintaining an equal number of positive and negative valence sentences with each group having an equal number of object and person subject nouns.

Creating the Yoked Design in the Read Condition

Data for the read condition were gathered after completion of the generate condition. Each participant in the generate condition was randomly paired with a participant in the read condition. We then created a set of complete sentences for the participant in the read condition using the generated words from the paired participant in the generate condition. To account for any generation errors, we created a set of contingencies. Should the participant in the generate condition have made a simple spelling error and the word was still reasonably recognizable, it would be corrected in the read condition. Should the participant in the generate condition have generated a word that was unrecognizable or completely failed to generate a word, the read condition sentence was completed using the most commonly generated word for that sentence. Both cases together accounted for less than 1% of all sentences. The yoked design accounted for any word effects where some words may be more memorable or interesting than other words. Doing this, we ensured that the target words were almost identical across both the generate and read conditions.

Counterbalance

The final 48 sentences were further divided into two lists of 24 sentences maintaining an equal number of positive and negative valence sentences as well as an equal number of object and person subject nouns within each list. Two lists were created because the self dimension was presented in blocked format. For half of all participants, List A was presented in the self condition while List B was presented in the other condition. Sentences within the self and other blocks were presented in a random order. To control for order and practice effects, both the list order (A or B) and the condition order (self or other) were counterbalanced. This created four possible sequences. For example, one sequence might begin with List A–self, followed by List A–other, then List B–self, and so on, while another sequence would start with the opposite list and condition.

Learning phase

The encoding paradigm we chose for the generate condition was low constraint open ended sentence completion. When participants generated the last word for each sentence, the only limitations were that they could only generate one word per sentence and they could not repeat any words. Prior to starting the learning phase, participants were informed of the later memory tests. Participants were also required to complete a practice trial before starting each block. Each trial consisted of one sentence, presented for 15 seconds that served as the time limit for participants to type out their generated word. In the read condition, participants were still required to physically type out a word. However, they were asked to copy the underlined word rather than to generate a word. This was done to control for any memory benefits conferred through motor movements during the performance of an action. After 15 seconds, the experiment would automatically advance to the next trial. There was no opportunity for participants to go back to previous trials

Distractor task

Two distractor tasks, one between the two blocks and the other after the completion of the learning phase, were used. Distractor tasks consisted of 10 basic two operation linear arithmetic problems (e.g., 2 + 5 - 10). They were presented one at a time for 12 seconds each. Participants were made to believe that we were collecting their performance data and were thus more encouraged to do their best.

Measures

Memory performance was measured with both free recall and cued recall memory tests. In the free recall test, participants were encouraged to recall as many target words as they could remember. There was no time limit. Following the free recall test, participants completed a cued recall test. Cues consisted of the presentation of the subject noun and verb (e.g., "Your mother made your ______" would turn into the cue "Mother made"). Participants were instructed to recall the target word they remembered generating or typing out. These cues were presented one at a time but there was no time limit.

Scoring

Performance was scored by the research team. Score was determined from a binary correct/incorrect scheme based on whether the words remembered during the free and cued recall matched the words from the learning phase. To be counted as correct, the words had to match perfectly. Only capitalization was disregarded. Spelling or pluralization mistakes were counted as incorrect. Should a participant generate a word with incorrect spelling, the recalled word in the testing phase had to match this incorrect spelling. Should a participant write multiple words in the recall tests such as including the possessive pronouns or adjectives, only the noun was scored. Should the participant break one of the generation limitations by either generating more than one word or generating a repeated word, those words were always counted as incorrect. Finally, should the participant fail to generate a word in the learning phase, it would always be counted as incorrect. These extraneous situations accounted for less than 1% of all sentences. Finally, scores were separated within subjects by self-condition and type of recall test (self-cued, self-free, other-cued, other-free) each providing a score out of 24.

Results

Demographics

A total of 89 undergraduate students participated in the study (68 female, 21 male). The mean age was 18.02 years (SD = 0.92). All participants had at least a high school diploma.

Statistical Analysis

All analyses were conducted in R using the afex, effsize, dplyr, rstatix, readxl, Rmisc, and ggplot2 packages. A 2 × 2 mixed ANOVA was performed using the aov_ez() function from the afex package, with generation (generate vs. read) as a between-subjects factor and self condition (self vs. other) as a within-subjects factor.

Assumption checks were conducted using Q–Q plots for normality and Bartlett's test for homogeneity of variance. Outliers were detected using the interquartile range (IQR) method within each condition.

Post hoc comparisons were conducted using Bonferroni-corrected t-tests via the pairwise.t.test() function. Descriptive statistics (means and standard errors) were calculated using the summarySEwithin() function from the Rmisc package. Interaction plots and figures were generated with ggplot2.

Outliers

To identify and remove potential outliers prior to analysis, we applied the interquartile range (IQR) method, a commonly used non-parametric approach based on the spread of the middle 50% of the data. Specifically, for each of the four experimental conditions (generate–self, generate–other, read–self, read–other), we first calculated the first quartile (Q1) and third quartile (Q3) of the score distributions. A data point was classified as an outlier if it fell below $Q1 - 1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$, which flags values that are substantially lower or higher than the typical range. This method was applied separately within each condition to preserve the integrity of condition-specific variability. Outliers identified by this method were subsequently removed from the dataset before conducting the primary analyses. In the cued recall scores, a total of 4 participants had scores meeting this criterion (2 generate, 2 read) for a final sample size of 85 (43 generate, 42 read). In the free recall scores, only 1 participant from the read condition met this criterion with a final sample size of 88 (45 generate, 43 read).

Assumption Checks

Prior to analysis, key assumptions were examined. Q–Q plots for both free and cued recall scores in each condition indicated that residuals were approximately normally distributed. (See Figure 3A and 3B in the Appendix). Bartlett's tests of homogeneity of variance were non-significant for the cued recall dataset ($ps > .90$), indicating that variances were equal across groups. However, for the free recall data, Bartlett's test for the generate condition was significant, $K^2(1) = 26.24$, $p < .001$, suggesting heterogeneity of variance. Given that the group sizes were nearly equal and that the design was simple, the mixed ANOVA was deemed robust to this violation, but results should be interpreted with caution.

Data Visualization and Confidence Intervals

To further visualize condition means and assess the magnitude of within-subject effects, confidence intervals were computed using the method described by Cousineau (2005), with the correction proposed by Morey (2008). This procedure adjusts for the underestimation of variability that arises when data are normalized within subjects to remove between-subject variability.

Normalization was performed separately within each level of the between-subjects factor (generate condition), as recommended for mixed designs. This ensured that within-subject comparisons (i.e., self vs. other) were calculated relative to each participant's own mean, while preserving the integrity of between-group structure. Confidence intervals were then derived from the normalized values and corrected using the factor $\sqrt{(n / (n - 1))}$, where n represents the number of within-subject conditions. All computations were performed using the summarySEwithin() function from the Rmisc package in R.

Cued Recall

A 2 (generation: generate vs. read; between-subjects) × 2 (self: self vs. other; within-subjects) mixed ANOVA was conducted on cued recall scores. The resulting analysis showed a signifciant main effect of generation ($F(1, 83) = 455.06$, $p < .001$, ges = .806), with the generation condition outperforming the self condition. It also showed a significant main effect of self ($F(1, 83) = 16.33$, $p < .001$, ges = .046), with the self condition outperforming the other condition. However, the generation × self interaction was not significant ($F(1, 83) = 0.67$, $p = .415$, ges = .002.)

Post Hoc Comparisons

Further post hoc comparisons were conducted to further analyze the main effects. In all scenarios, the Bonferroni correction was applied. Within the generate dimension, paired t-tests revealed a significant self-reference effect with the self condition outperforming the other condition for both the generate ($p = .043$, Cohen'd = .33) and the read conditions ($p < .001$, Cohen's d = .52). Within the self dimension, paired t-tests revealed a significant generation effect with the generate condition outperforming the read condition for both the self ($p < .001$, Cohen's d = 3.73) and other conditions ($p < .001$, Cohen's d = 4.38). Visualization of cued recall scores, as well as the results of post hoc comparisons, can be found in Figure 1A.

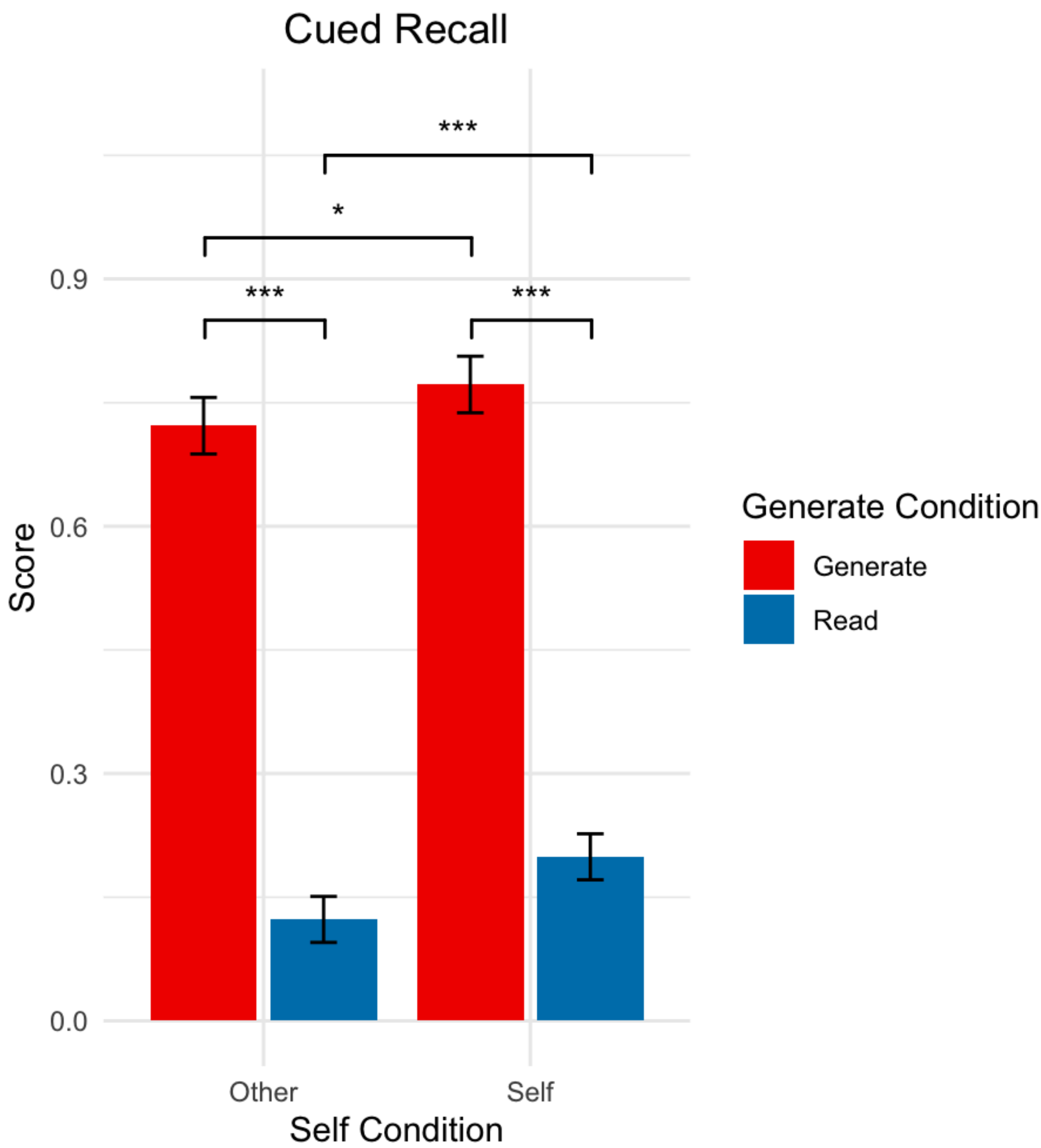
Cued Recall
***
*
***
***
0.9
0.6
0.3
0.0
Score
Other
Self
Self Condition
Generate Condition
Generate
Read

Figure 1A

Mean cued recall scores as a function of self condition (self vs. other) and generate condition (generate vs. read). Error bars represent within subjects corrected 95% confidence intervals (Morey, 2008). Asterisks indicate significance levels: p < .05 (*), p < .01 (**), p < .001 (***)

Free Recall

A 2 × 2 mixed ANOVA with conditions identical to the one performed on cued recall scores was performed on free recall scores. The analysis showed a signficiant main effect of generation ($F(1, 86) = 31.54$, $p < .001$, ges = .214.), with the generate condition outperforming the read condition. It also showed a significant main effect of self ($F(1, 86) = 7.96$, $p = .006$, ges = .023), with the self condition outperforming the other condition. The generation × self interaction was again non-significant, ($F(1, 86) = 0.43$, $p = .515$, ges = .001.)

Post Hoc Comparisons

As we did with cued recall, post hoc comparisons were made to further analyze the main effects. In all scenarios, the Bonferroni correction was applied. Within the generate dimension, paired t-tests revealed a significant self-reference effect with the self condition outperforming the other condition for both the generate ($p = .048$, Cohen'd = .307) and the read conditions ($p = .036$. Cohen'd = 0.338).

Due to the significant violation of the homogeneity of variance assumption in the free recall data (Bartlett's $K^2 = 26.24$, $p < .001$), follow-up Welch's t-tests were instead conducted within the self dimension to confirm the reliability of the between-subjects generate condition comparisons. Paired t-tests revealed a significant generation effect with the generate condition outperforming the read condition for both the self ($p < .001$, Cohen's d = 1.01) and other conditions (p < .001, Cohen's d = 1.06). Visualization of cued recall scores, as well as the results of post hoc comparisons, can be found in Figure 1B.

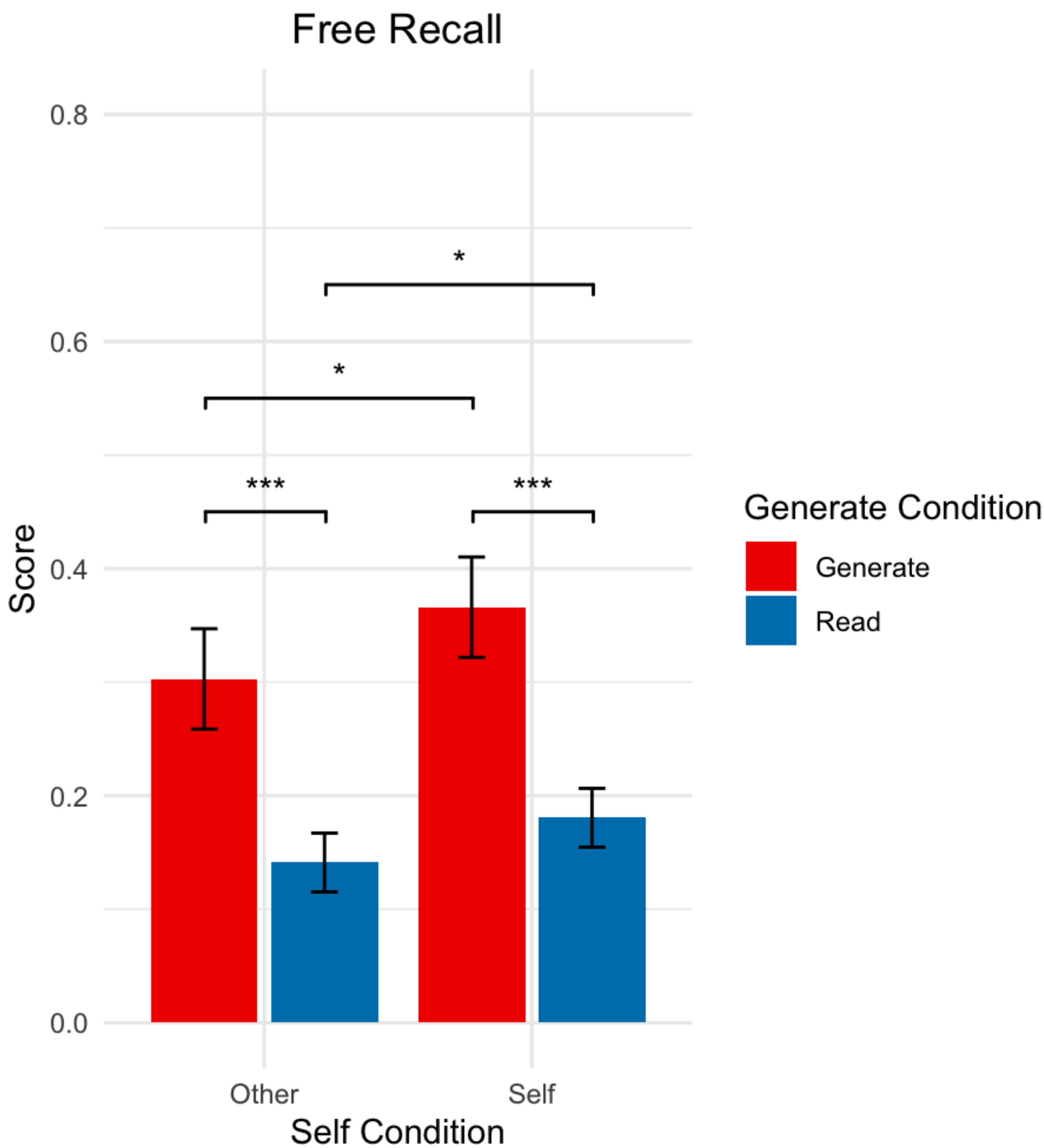

Figure 1B

Mean free recall scores as a function of self condition (self vs. other) and generate condition (generate vs. read). Error bars represent within subjects corrected 95% confidence intervals (Morey, 2008). Asterisks indicate significance levels: $p < .05$ (*), $p < .01$ (**), $p < .001$ (***)

Summary

Across both cued and free recall tasks, participants remembered more when they generated words and when sentences referred to the self. However, the lack of interaction suggests that the generation and self-reference effects were additive, not interactive, in this paradigm.

Recall Condition

As an exploratory analysis, we examined the effect of recall conditions (cued vs free) on the previously examined conditions. To investigate this, a 2x2x2 mixed ANOVA was conducted with conditions generation (generate vs. read; between-subjects), self-reference (self vs. other; within-subjects), and recall type (free vs. cued; within-subjects).

For main effects, the analysis revealed a main effect of generation ($F(1, 87) = 165.86$, $p < .001$, ges = .508), indicating that participants in the generate condition outperformed those in the read condition overall. A main effect of self-reference was also significant ($F(1, 87) = 18.70$, $p < .001$, ges = .026) indicating better memory performance for self-relevant items than for other-relevant ones. Finally, the main effect of recall type was significant, $F(1, 87) = 100.18$, $p < .001$, ges = .228, showing that cued recall produced better memory performance than free recall.

For interactions, the generation × self-reference interaction ($F(1, 87) = 0.03$, $p = .870$), the self-reference × recall type interaction ($F(1, 87) = 0.02$, $p = .890$), and the three-way interaction between generation, self-reference, and recall type ($F(1, 87) = 0.79$, $p = .376$.) were all found to be not significant. However, the generation × recall type interaction was significant ($F(1, 87) = 106.53$, $p < .001$, ges = .239), which suggests that the memory benefit of generation was greater for cued recall than for free recall.

To follow up the significant generation × recall type interaction, we conducted pairwise comparisons using estimated marginal means. Results showed that participants in the generate condition significantly outperformed those in the read condition on both cued recall, $t(87) = 16.39$, $p < .001$, and free recall, $t(87) = 4.77$, $p < .001$. However, the generation effect was markedly larger in the cued recall condition (mean difference = 0.55) than in free recall (mean difference = 0.16), consistent with the strong interaction observed in the ANOVA. A line graph was plotted to show this interaction in Figure 2.

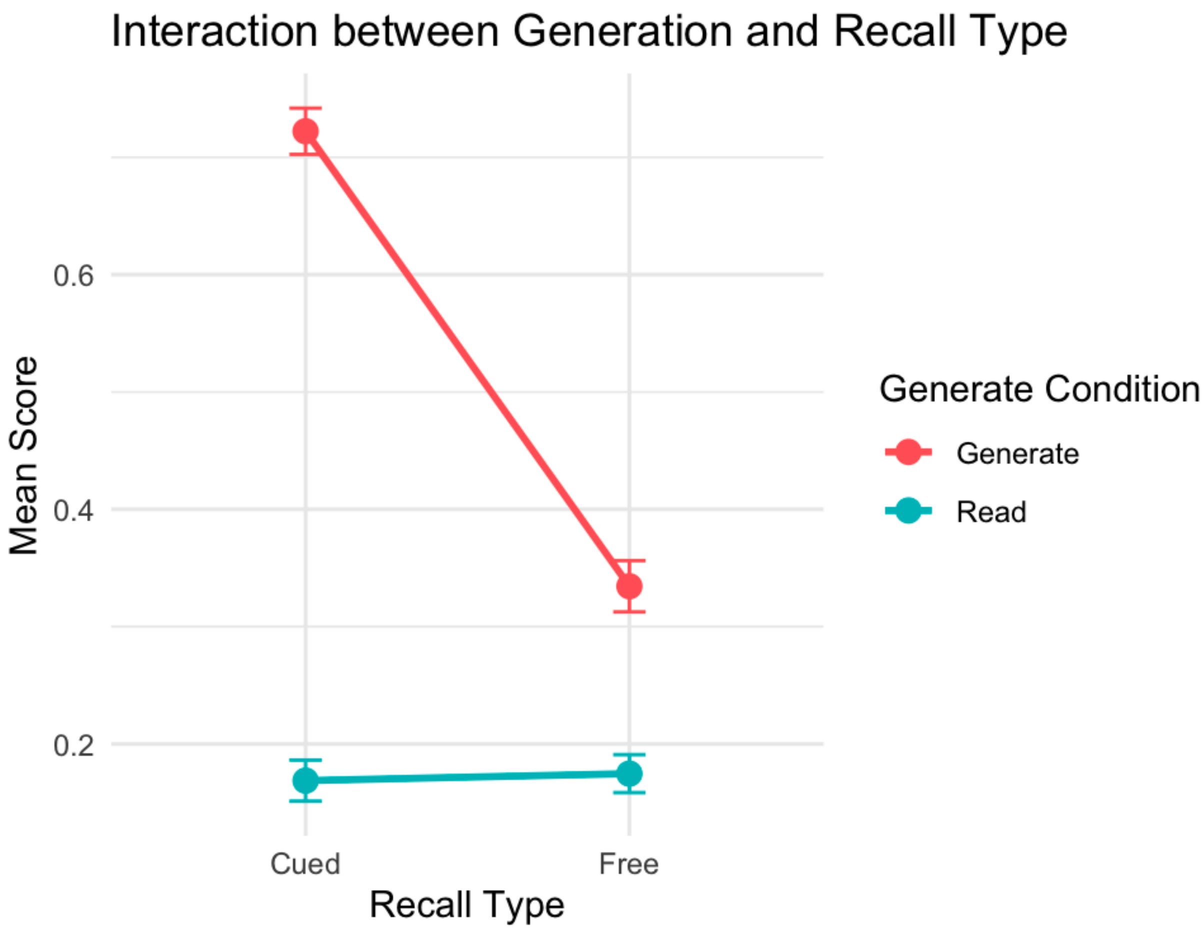

Figure 2

Interaction between generation condition and recall type on memory performance. Error bars represent standard error.

Discussion

Generation Effect

Across both the cued and free recall tasks, the generation condition showed a large improvement in memory over the read condition. Participants in the generate condition significantly outperformed those in the read condition, regardless of testing condition (cued vs free) and self condition. This replicates prior findings on the generation effect in which actively generating information enhances memory encoding compared to passive reading (Slamecka & Graf, 1978). Notably, the effect size for generation was substantial in both tasks (ges = .806 and .214 for cued and free recall, respectively), with all post hoc comparisons yielding highly significant differences. These results provide further support to the extensive literature on the robustness and generalizability of generation-based encoding strategies, highlighting the benefits of active learning through deeper semantic processing, enhanced item distinctiveness, and stronger memory traces. From a practical standpoint, the findings reinforce the utility of learner-active techniques in educational and cognitive training settings, where encouraging generation may lead to more durable learning outcomes.

Interactions

One goal of the present study was to test whether the generation and self-reference effects followed an (1) additive model or (2) an interactive model. While we predicted an interactive model, our testing paradigm did not result in a significant interaction. This suggests that the relationship between generation and self-reference may be orthogonal and therefore additive. That is, the self-reference effect is not affected by the presence of a generative activity and vice versa. This likely indicates that the two effects operated independently at least in the context of how they were manipulated in this study.

From a multi-factor perspective, it is likely that the cognitive mechanisms underlying the generation and self-reference effects are distinct, which may explain the absence of an interaction in the present study. The generation effect is often attributed to processes such as effortful retrieval, enhanced semantic elaboration, and the production effect, all of which increase the distinctiveness of the generated item in memory (Craik & Tulving, 1975; McCurdy et al., 2020). When individuals generate information rather than passively read it, they are more likely to engage in active meaning construction, which facilitates deeper encoding and promotes long-term retention.

In contrast, the self-reference effect is thought to operate primarily through self-schema activation. By encoding information in relation to the self, participants are likely to form more personally meaningful representations, drawing on autobiographical knowledge and emotional salience (Sui & Humphreys, 2015; Cunningham et al., 2024). These features enhance memory by strengthening the associative links between the new information and the self.

It is likely that these two effects rely on different encoding pathways, with generation driven by cognitive factors like production and elaboration and self-reference by affective and self-referential relations. While both improve memory, they contribute independently without interacting. This interpretation aligns with multifactorial models of memory (Hirshman & Bjork, 1988; McCurdy & Leshikar, 2022), allowing for multiple memory enhancement factors that can contribute in parallel, particularly when they are not functionally redundant.

To better understand the mechanisms underlying the observed memory effects, it is helpful to consider how performance varies across different testing formats. While free and cued recall both rely on relational memory processes, it is to differing extents. Cued recall is often more constrained by cue-target associations, while free recall may draw more heavily on contextual relational memory. In this study, we used both recall types to evaluate how the generation and self-reference effects operate within relational memory encoding. The results revealed a strong interaction between generation and recall type, with the generation effect being larger in the cued recall condition. This suggests that generation particularly benefits cue-dependent associative retrieval, possibly by strengthening links between studied items and their retrieval cues. In contrast, generation conferred only modest benefits in free recall, where the absence of cues requires greater reliance on broader contextual memory. This aligns with previous findings of larger generation effects in cued over free recall conditions (Hirshman & Bjork, 1988; McCurdy et al. 2020). Meanwhile, the self-reference effect produced smaller, consistent gains across both formats and did not interact significantly with recall type. This suggests that it may influence encoding through factors that are less sensitive to the recall format, which contrasts previous findings where the self-reference effect shows greater effect sizes in cued recall conditions when compared to free recall conditions (Symons & Johnson, 1997). However, this discrepancy may be due to our testing paradigm as their methods relied on relatedness judgements whereas we used self-referential personal pronouns. Despite these differences, these findings support the idea that generation and our version of self-reference may engage distinct encoding processes. Future studies could more clearly differentiate these mechanisms by incorporating recognition tasks that emphasize item-specific memory.

Self-Referential Possessive Pronouns

The second goal of our experiment was to explore a potential testing paradigm using second person possessive pronouns (your) to induce the self-reference effect. Our experiment found a small but significant main effect of using the self-referential "your" as opposed to the other-referential "their" in sentence completion tasks. This effect persists regardless of generate condition or type of recall test. With these results, we provide support to previous literature which found that possession confers memory advantages (Sui & Humphreys, 2013). These findings also mirror the study done by Ditman et al. (2010) in which first person pronouns were used to enhance memory. To our knowledge, we are the first study to employ a language based paradigm with possession, in which the self condition manipulation is expressed subtly and solely through wording differences. Considering the results, this ultimately suggests that the simple change of using the word “your” instead of “their” during encoding increases subsequent memory outcomes significantly.

One potential explanation for this benefit may be the described self-reference effect. In alignment with the depth-of-processing framework (Craik & Tulvig, 1975), the self is processed on a deeper level. Because the self acts as a highly elaborated and frequently accessed cognitive structure, it offers a rich associative network that facilitates encoding through multiple meaningful connections to self-related schema. When participants read a sentence like “Your glass contained your _____”, it may invoke a sense of familiarity and engage said self-related schema. On the other hand, a sentence like “Their glass contained their _____” is more detached or observational and likely lacks self-relevance. This aligns with Rogers et al. (1977), who originally demonstrated that self-referential encoding leads to superior recall compared to semantic or structural encoding, even when participants are not explicitly instructed to remember the material.

Further theoretical accounts, such as Symons and Johnson’s (1997) review, suggest that the self-reference effect may arise from both enhanced elaboration and increased organizational processing. When information is linked to the self, individuals are more likely to organize it around a central, familiar theme, thereby promoting better integration and retrieval. Our study's use of second-person pronouns may have prompted this kind of processing implicitly. While participants may not have consciously related the material to themselves, the usage of “your” was sufficient to activate self-related networks.

Another potential explanation for these effects may be the increased ease of mental imagery associated with self-referential sentences. Prior research suggests that the ability to form vivid mental images during encoding can enhance memory by creating more elaborated and interconnected representations (Nilsson, 2000; Paivio, 1971; Zimmer et al., 2001;). In the case of self-referential language, such as “your mother” or “your glass”, participants are likely to generate more concrete and personalized mental images, drawing on their own experiences and associations (Bower& Gilligan, 1979). In contrast, phrases like “their mother” or “their glass” are less personally grounded, leading to more generic imagery or even a lack of imagery.

This idea is supported by the dual-coding theory (Paivio, 1971), which posits that information encoded both verbally and visually has a higher likelihood of being remembered. When the self is invoked, the imagery tends to be more vivid, emotionally salient, and semantically rich, providing both visual and conceptual codes that facilitate later recall. Moreover, the elaboration hypothesis (Klein & Kihlstrom, 1986) suggests that self-referential processing leads to deeper encoding by encouraging elaboration through the integration of personal experiences. “Your” sentences may be more likely to involve these personal experiences and thus were better remembered. Taken together, these theories suggest that the memory advantage observed for “your” sentences may stem not only from deeper conceptual processing but also from enhanced imageability.

In sum, these findings offer support for the idea that minimal self-cueing, even without explicit instructions, can trigger deeper level encoding seen in other studies of the self-reference effect. This has broader implications for how linguistic framing might be used to boost memory in educational or real-world settings by subtly invoking the self without demanding active self-reflection.

Practical Implications

The present findings offer several important implications for both theoretical understanding and applied contexts. Most obviously, the consistent and robust generation effect across both cued and free recall conditions suggests that encouraging individuals to actively generate information can significantly improve memory over passive reading. This supports a long body of evidence favoring generative learning strategies in educational and instructional settings. Educators and curriculum designers may benefit from incorporating sentence completion tasks, prediction exercises, or retrieval-based study methods that allow learners to construct or infer content on their own. These methods are particularly promising for improving long-term retention in both classroom and digital learning environments.

The study also demonstrated a significant but smaller self-reference effect induced by minimal linguistic manipulation through the use of second-person possessive pronouns ("your") compared to third-person ("their"). This finding suggests that subtle personalization of material, even without overt self-reflection or instructions, is sufficient to elicit measurable improvements in memory. Instructional materials and training modules might benefit from framing content in the second person. For instance, reframing "The brain controls the body" as "Your brain controls your body" can prompt individuals to encode information more deeply by activating self-relevant schemas or enhancing imageability.

The study found no significant interaction between generation and self-reference effects, indicating that the benefits of each strategy are additive. In practical terms, this means that using both strategies together does not diminish their individual effectiveness. Learners who are asked to complete sentences that are self-referential can benefit from both the generative effort and the self-related encoding independently. When considering instructional design, combining personalized content with active engagement techniques can result in cumulative memory benefits. For example, presenting a sentence like "Your brain controls your..." while asking students to fill in the rest of the sentence may be a learning task that simultaneously taps into both mechanisms.

Finally, the finding that generation effects were especially pronounced under cued recall conditions supports the idea that generative strategies are particularly effective when learners are expected to retrieve specific associations. This suggests a practical advantage in structured learning environments where cues or prompts are provided. Conversely, the smaller and more stable self-reference effect suggests it may operate more generally and could be used across different retrieval contexts without concern for format.

In sum, these findings support a multi-factor approach to improving memory. Educators may enhance learning by encouraging learners to actively generate content and by subtly linking material to the self. As there is no interaction, these strategies can be layered for greater effect.

Limitations and Future Directions

While we are confident in our results, several limitations should be acknowledged. First, the sample was largely homogeneous, consisting primarily of undergraduate students with a narrow age range. This limits the generalizability of the findings to more diverse populations. Future studies may benefit from studying alternative student populations such as those in grade school as the self-reference effect may differ between adults and adolescents (Leshikar et al., 2015). In addition, grade school populations likely serve to gain the most benefit from learning memory enhancement techniques.

One notable limitation of the current study concerns the nature of the self-reference manipulation. Our self-reference paradigm of using second-person possessive pronouns ("your") to induce self-relevance differed from the traditional methods involving explicit self-referential judgments, such as asking participants whether a word describes themselves. While this novel, language-based manipulation successfully elicited a measurable self-reference effect, it may not engage the same depth of self-referential processing typically observed in studies using explicit relatedness judgments (e.g., Rogers et al., 1977). As a result, the cognitive mechanisms underlying the observed effect may differ from those in more established self-reference literature. It is possible that our experiment did not find a significant interaction because of this difference. Thus, caution is warranted when interpreting the generalizability of our findings.

Additionally, the study did not find a significant interaction between self-reference and recall type. Prior research using traditional self-reference paradigms has shown that the benefits of self-referential encoding tend to be more pronounced in cued recall than in free recall, likely due to enhanced cue-target relational processing when participants make relatedness judgements (Symons & Johnson, 1997). The absence of such an interaction in the present study may stem from the reduced depth of processing, which may not have been sufficient to differentially strengthen cue-based memory retrieval.

Future research should consider incorporating a comparison between implicit and explicit self-reference manipulations to better understand the conditions under which self-relevance enhances memory. This can be potentially achieved by asking participants to make judgements of relatedness after the generation/reading of each sentence. This would allow us to examine if there is a correlation between second person possessive pronouns and judgements of relatedness. Additionally, examining whether deeper self-referential processing produces larger effects in cued recall relative to free recall would help solidify the role of encoding depth and relational memory processes as factors in the self-referential effect.

Another limitation of the present study is the absence of a dedicated retrieval manipulation. While the design focused on encoding strategies, it did not explore how these effects interact with different retrieval conditions, such as recognition tests or retrieval practice. Future studies could include a retrieval-focused component or a recognition memory task. This would be useful to examine the multiple factors involved and help disentangle whether these two effects operate primarily through relational or item specific memory.

Finally, there may have been merit in including a condition in which first person is used (e.g., My mother made my _____). Because generation is done by the participant, it may cause them to reframe the sentence to their own perspective. This would cause "your" to no longer refer to the participant and therefore, "my" would be more appropriate to induce a self reference effect. Thus, we may have needed to use first person possessive pronouns in our generate conditions in order to see an interaction. A future study should include a first person condition in order to test this hypothesis.

Conclusions

The present study explored the individual and combined effects of generation and self-reference on memory performance using sentence completion tasks. Across both free and cued recall, the generation effect produced robust and consistent improvements in memory, reinforcing prior findings that actively producing information enhances encoding and retrieval. Additionally, our novel self-reference manipulation of using second-person possessive pronouns showed a small but significant memory advantage, supporting the idea that even subtle linguistic cues can activate self-related processing and enhance retention.

Importantly, no significant interaction was found between generation and self-reference, suggesting that these strategies operate additively rather than redundantly. This implies that learners can benefit from employing both strategies simultaneously without risk of interference, expanding their practical utility in educational contexts. Furthermore, the lack of a three-way interaction with recall type indicates that the effects of generation and self-reference generalize across different forms of relational memory tests.

Together, these findings contribute to our understanding of how distinct encoding strategies influence memory and highlight the potential of linguistically grounded self-reference manipulations. Future work should further investigate the relation between our personal pronouns research paradigm and traditional self-reference methods. It may also be fruitful to explore its interactions between traditional self-reference methods and generation in both recall and retrieval contexts. Ultimately, identifying effective strategies for enhancing memory remains an important goal for both cognitive research and applied learning environments.

References

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology, 63*(4), 596. https://doi.org/10.1037/0022-3514.63.4.596

Banaji, M. R., & Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology, 45*(1), 297–332.

Berrin-Wasserman, S., Winnick, W. A., & Borod, J. C. (2003). Effects of stimulus emotionality and sentence generation on memory for words in adults with unilateral brain damage. *Neuropsychology, 17*(3), 429–438. https://doi.org/10.1037/0894-4105.17.3.429

Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. *Journal of Research in Personality, 13(*4), 420–432. https://doi.org/10.1016/0092-6566(79)90005-9

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*, 145–153. https://doi.org/10.1016/0361-476X(79)90069-9

Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science, 20(*1), 27–32. https://doi.org/10.1111/j.1467-9280.2008.02249

Bugg, J. M., & McDaniel, M. A. (2012). Selective benefits of question self-generation and answering for remembering expository text. *Journal of Educational Psychology, 104*(4), 922-931. https://doi.org/10.1037/a0028661

Carroll, M., Davis, R., & Conway, M. (2001). The effects of self-reference on recognition and source attribution. *Australian Journal of Psychology, 53*(3), 140–145. https://doi.org/10.1080/00049530108255136

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences, 6*(4), 347–365. https://doi.org/10.1207/s15327809jls0604_1

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods For Psychology, 1*(1), 42–45.

Cowan, N., Bao, C., Bishop-Chrzanowski, B. M., Costa, A. N., Greene, N. R., Guitard, D., Li, C., Musich, M. L., & Ünal, Z. E. (2024). The Relation Between Attention and Memory. *Annual review of psychology, 75*, 183–214. https://doi.org/10.1146/annurev-psych-040723-012736

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Cunningham, S. J., Ahmed, Z., March, J., Golden, K., Wilks, C., Ross, J., & McLean, J. F. (2024). Put you in the problem: Effects of self-pronouns on mathematical problem-solving. *Quarterly Journal of Experimental Psychology, 77*(2), 308–325. https://doi.org/10.1177/17470218231174229

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2014). The self-reference effect on memory in early childhood. *Child Development, 85*(2), 808–823. https://doi.org/10.1111/cdev.12144

D'Argembeau, A., Comblain, C., & Van der Linden, M. (2005). Affective valence and the self-reference effect: Influence of retrieval conditions. *British Journal of Psychology, 96*, 457-66. https://doi.org/10.1348/000712605X53218

Ditman, T., Brunyé, T. T., Mahoney, C. R., & Taylor, H. A. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition, 115*(1), 172–178. https://doi.org/10.1016/j.cognition.2009.10.014

Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Education and Psychology, 70*(2), 109–118. https://doi.org/10.1037/0022-0663.70.2.109

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage, 62*(1), 562-574.

Geghman, K. D., & Multhaup, K. S. (2004). How generation affects source memory. *Memory & Cognition, 32*(5), 819–823. https://doi.org/10.3758/BF03195871

Ghatala, E. S. (1981). The effect of internal generation of information on memory performance. *The American Journal of Psychology*, 443–450. https://doi.org/10.2307/1422255

Guskjolen, A., Cembrowski, M.S. Engram neurons: Encoding, consolidation, retrieval, and forgetting of memory. *Mol Psychiatry, 28*, 3207–3219 (2023). https://doi.org/10.1038/s41380-023-02137-5

He, L., Han, W., & Shi, Z. (2021). Self-reference effect induced by self-cues presented during retrieval. *Frontiers in Psychology, 12*, 562359. https://doi.org/10.3389/fpsyg.2021.562359

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 484–494. https://doi.org/10.1037/0278-7393.14.3.484

Holley, C. D., Dansereau, D. F., McDonald, B. A., Garland, J. C., & Collins, K. W. (1979). Evaluation of a hierarchical mapping technique as an aid to prose processing. *Contemporary Educational Psychology, 4*, 227–237. https://doi.org/10.1016/0361-476X(79)90043-2

Hood, B. (2012). The Self Illusion: Why there is no 'you' inside your head. Hachette UK.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20*(5), 497–514. https://doi.org/10.1016/S0022-5371(81)90138-9

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32*(4), 421–445. https://doi.org/10.1006/jmla.1993.1023

Hunt, R. R. (2003). Two Contributions of Distinctive Processing to Accurate Memory. *Journal of Memory and Language, 48*, 811-825. https://doi.org/10.1016/S0749-596X(03)00018-4

Kane, J. H., & Anderson, R. C. (1978). Depth of processing and interference effects in the learning and remembering of sentences. *Journal of Educational Psychology, 70*(4), 626–635. https://doi.org/10.1037/0022-0663.70.4.626

Klein, S. B., & Kihlstrom, J. F. (1986). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General, 115*(1), 26–38. https://doi.org/10.1037/0096-3445.115.1.26

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology, 55*(1), 5–11. https://doi.org/10.1037/0022-3514.55.1.5

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science, 4(*1), 2515245920951503. https://doi.org/10.1177/2515245920951503

Leopold, C., & Mayer, R. E. (2015). An imagination effect in learning from scientific text. *Journal of Education and Psychology, 107,* 47–63. https://doi.org/10.1037/a0037142

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging, Neuropsychology, and Cognition, 22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Leutner, D., & Schmeck, A. (2014). The drawing principle in multimedia learning. In R. E. Mayer (Ed.), The Cambridge handbook of multimedia learning (Second Edition ed., pp. 433–448). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: effects of drawing and mental imagining text content. Computers in *Human Behavior, 25,* 284–289. https://doi.org/10.1016/j.chb.2008.12.010

Marsh, E. J. (2006). When does generation enhance memory for location? Journal of *Experimental Psychology: Learning, Memory, and Cognition, 32*(5), 1216–1220. https://doi.org/10.1037/0278-7393.32.5.1216

McCurdy, M. P., & Leshikar, E. D. (2022). Contextual framework of the generation effect. *The American Journal of Psychology, 135*(3), 251–270. https://doi.org/10.5406/19398298.135.3.01

McCurdy, M. P., Leach, R. C., & Leshikar, E. D. (2017). The generation effect revisited: Fewer generation constraints enhances item and context memory. *Journal of Memory and Language, 92*, 202–216. https://doi.org/10.1016/j.jml.2016.06.007

McCurdy, M. P., Viechtbauer, W., Sklenar, A. M., Frankenstein, A. N., & Leshikar, E. D. (2020). Theories of the generation effect and the impact of generation constraint: A meta-analytic review. *Psychonomic Bulletin & Review, 27*, 1139–1165. https://doi.org/10.3758/s13423-020-01762-3

McDaniel, M. A., Riegler, G. L., & Waddill, P. J. (1990). Generation effects in free recall: Further support for a three-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(5), 789–798. https://doi.org/10.1037/0278-7393.16.5.789

McFarland Jr, C. E., Frey, T. J., & Rhodes, D. D. (1980). Retrieval of internally versus externally generated words in episodic memory. *Journal of Verbal Learning and Verbal Behavior, 19*(2), 210–225. https://doi.org/10.1016/S0022-5371(80)90182-6

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*(2), 61–64.

Mulligan, N. W. (2011). Generation disrupts memory for intrinsic context but not extrinsic context. *The Quarterly Journal of Experimental Psychology, 64*(8), 1543–1562. https://doi.org/10.1080/17470218.2011.562980

Nilsson, L.-G. (2000). Remembering actions and words. In F. I. M. Craik & E. Tulving (Eds.), *Oxford handbook of memory* (pp. 137–148). Oxford:Oxford University Press.

Paivio, A. (1971). Imagery and language. In Imagery (pp. 7-32). Academic Press. https://doi.org/10.1016/B978-0-12-635450-8.50008-X

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology/Revue canadienne de psychologie, 45(3), 255–287. https://doi.org/10.1037/h0084295

Ponce, H. R., & Mayer, R. E. (2014). Qualitatively different cognitive processing during online reading primed by different study activities. *Computers in Human Behavior, 30*, 121–130. https://doi.org/10.1016/j.chb.2013.07.054

Qiu, X., & Jiang, F. K. (2021). Stance and engagement in 3MT presentations: How students communicate disciplinary knowledge to a wide audience. *Journal of English for Academic Purposes, 51*, 100976. https://doi.org/10.1016/j.jeap.2021.100976

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. Journal of personality and social psychology, 35(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Sedikides, C., & Green, J. D. (2000). On the self-protective nature of inconsistency-negativity management: using the person memory paradigm to examine self-referent memory. *Journal of Personality and Social Psychology, 79*(6), 906–921. https://doi.org/10.1037/0022-3514.79.6.906

Sedikides, C., & Green, J. D. (2004). What I don't recall can't hurt me: Information negativitiy verus information consistency as determinants of memorial self-defense. *Social Cognition, 22*(1), 4-29. https://doi.org/10.1521/soco.22.1.4.30987

Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Education and Psychology, 102,* 872–879. https://doi.org/10.1037/a0019640

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

Sui, J., & Humphreys, G. W. (2013). Self-referential processing is distinct from semantic elaboration: Evidence from long-term memory effects in a patient with amnesia and semantic impairments. *Neuropsychologia, 51*(13), 2663-2673. https://doi.org/10.1016/j.neuropsychologia.2013.07.025

Sui, J., & Humphreys, G. W. (2015). The integrative self: How self-reference integrates perception and memory. *Trends in Cognitive Sciences, 19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human perception and performance, 38*(5), 1105–1117. https://doi.org/10.1037/a0029792

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychological Bulletin, 121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Zimmer, H. D., Cohen, R. L., Guynn, M. J., Engelkamp, J., Kormi-Nouri, R., & Foley, M. A. (2001). *Memory for action*. New York: Oxford University

Appendix

Table 1

List of the subject nouns and verbs used to construct the sentence materials in the study, along with mean difficulty and valence ratings provided by independent raters (N = 7). Difficulty was rated on a 0–10 Likert scale, where 0 indicated "not difficult at all" and 10 indicated "extremely difficult." Valence was rated on a similar scale, where 0 indicated either "not negative at all" or "not positive at all" and 10 indicated either "very negative" or "very positive". The final set of items was selected to ensure a balanced distribution of valence (positive vs. negative) and subject type (person vs. object), while maintaining low overall difficulty to ensure readability and clarity for participants.

Subject Noun	Verb	Difficulty	Valence
Teacher	**Enjoyed**	0.142857	8.142857
Mother	**Made**	0.428571	7.142857
Technician	**Fixed**	0.142857	6.571429
Carpenter	**Installed**	2.285714	4.714286
Accountant	**Finished**	0.428571	5.857143
Gardener	**Trimmed**	0	5.714286
Painting	**Moved**	2.142857	6.285714
Money	**Motivated**	2.714286	5
Phone	**Located**	1.285714	7.142857
Present	**Surprised**	0	7.714286
Letter	**Comforted**	0.571429	9.428571
Camera	**Captured**	0	5.714286
Jar	**Contained**	0.428571	5.428571
Knife	**Cut**	0.285714	4.428571

Suitcase	**Carried**	0	4.714286
Umbrella	**Covered**	0.285714	6.714286
Ice	**Chilled**	0	5.571429
Towel	**Dried**	0	4.285714
Lawyer	**Defended**	0.142857	7.285714
Customer	**Thanked**	0.428571	6.571429
Nurse	**Calmed**	1.714286	4.714286
Horse	**Licked**	2.428571	4
Coach	**Encouraged**	0	6.857143
Child	**Helped**	3	7.428571
Chef	**Critiqued**	0.428571	4.285714
Baker	**Burned**	0	6.428571
Butcher	**Spoiled**	0.285714	7
Barista	**Spilled**	0	5.857143
Sailor	**Sank**	0.285714	8.714286
Potter	**Shattered**	2	7.142857
Boot	**Tripped**	0.857143	7.714286
Costume	**Scared**	1.285714	4.428571
Video	**Toddler**	0	4.285714
Article	**Misled**	0	7.428571
Virus	**Sickened**	0.428571	8.714286
Dance	**Disappointed**	0	7.714286
Garbage	**Stunk**	2	7.285714
Sand	**Clogged**	2.285714	6.142857
Mud	**Stained**	0.714286	5.428571
Pollution	**Poisoned**	0	9.714286

Ball	**Broke**	0.142857	8.571429
Water	**Damaged**	0.428571	7.428571
Criminal	**Escaped**	2.857143	9.285714
Instructor	**Failed**	1.571429	7.714286
Villian	**Killed**	2.428571	10
Spy	**Exposed**	1.857143	8.857143
Tyrant	**Enslaved**	2.714286	9.571429
Opposition	**Beat**	0.714286	7.285714

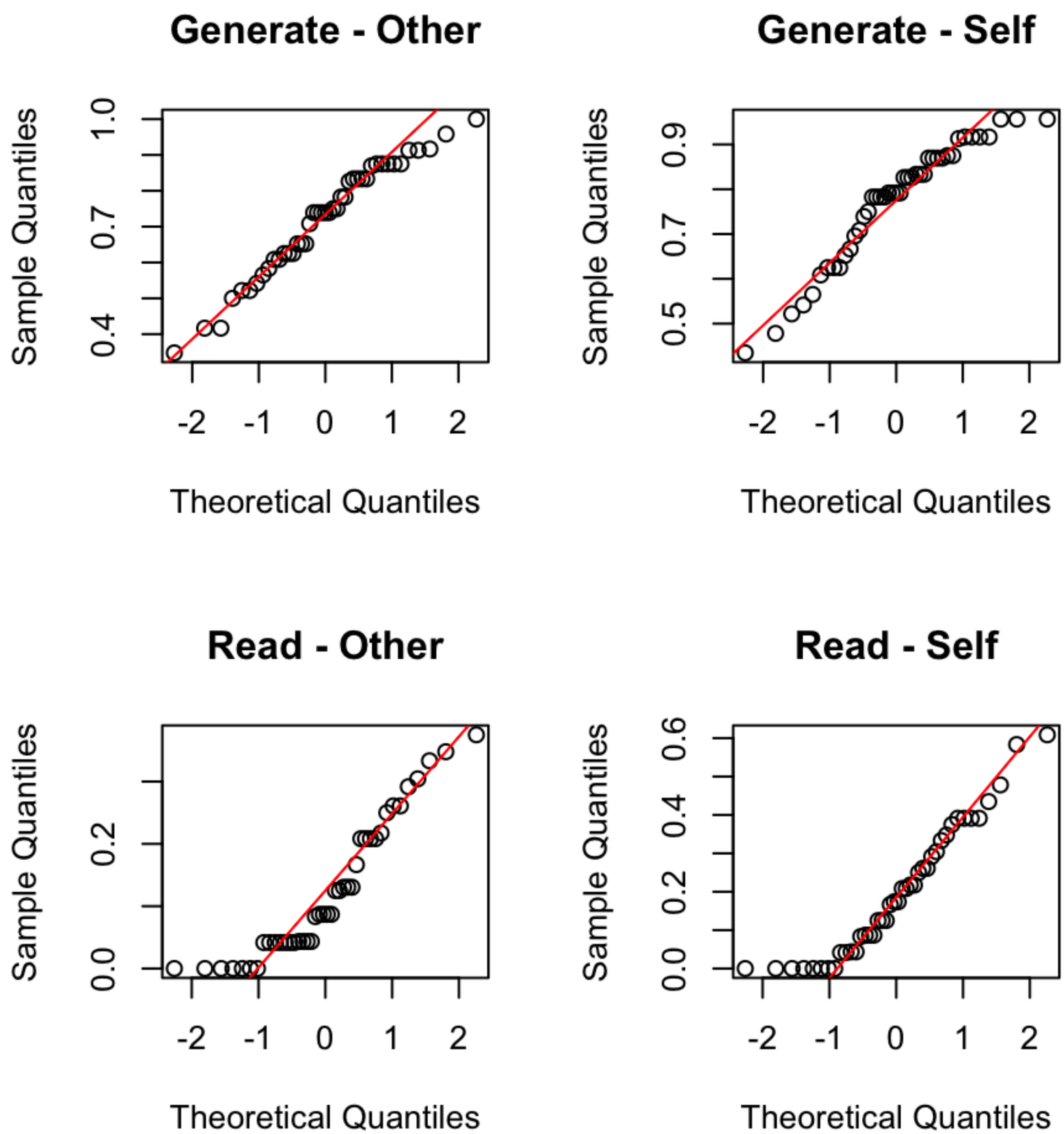

Figure 3A

Q–Q plots assessing the normality of residuals within the cued recall scores for each condition: generate–other, generate–self, read–other, and read–self. Each plot compares the sample quantiles of participants' recall scores to the theoretical quantiles of a normal distribution. All *conditions show reasonably linear patterns, suggesting that the assumption of normality was adequately met.*

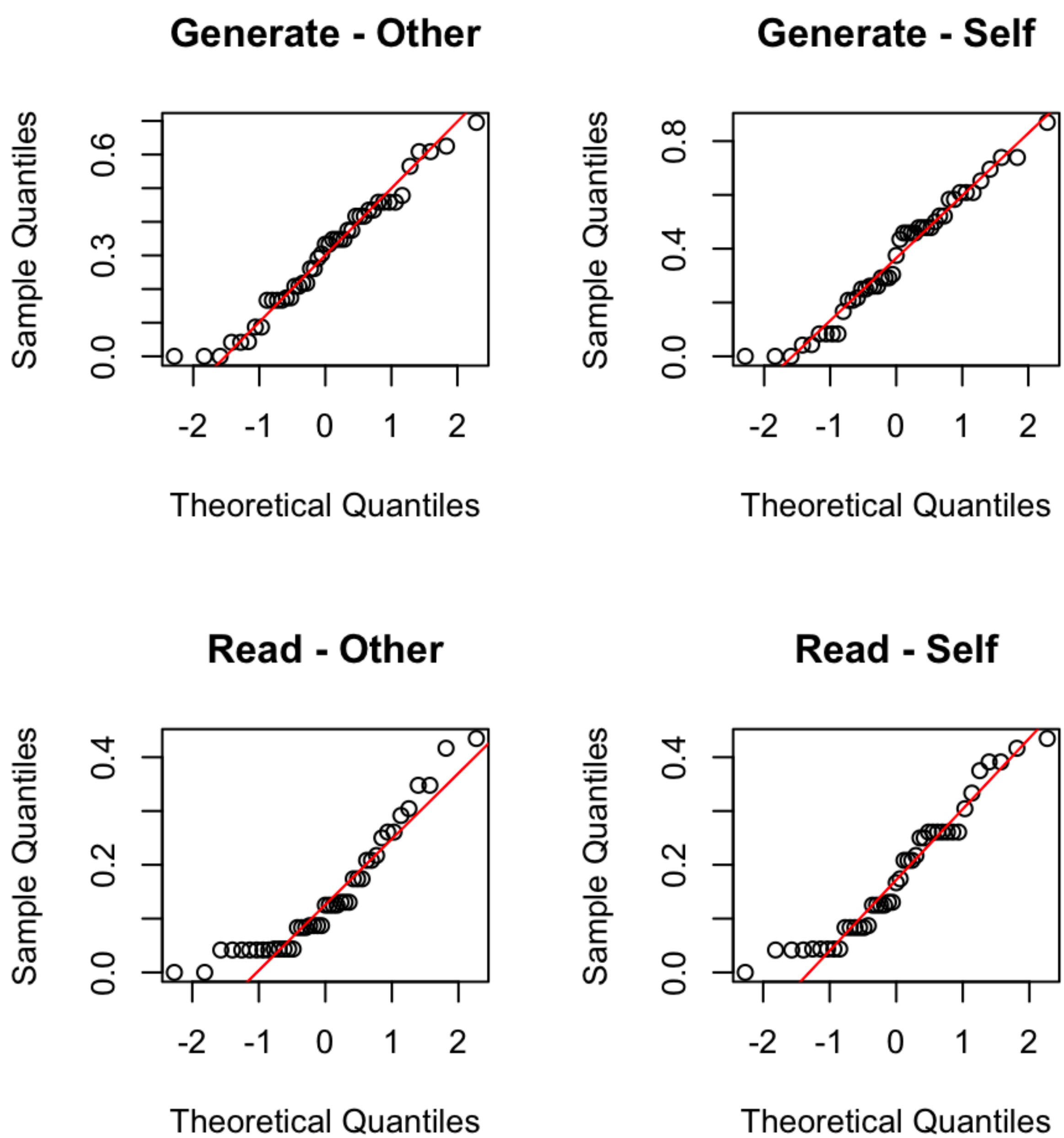

Figure 3B

Q–Q plots assessing the normality of residuals within the free recall scores for each condition: generate–other, generate–self, read–other, and read–self. Each plot compares the sample quantiles of participants' recall scores to the theoretical quantiles of a normal distribution. All conditions show reasonably linear patterns, suggesting that the assumption of normality was adequately met.