THE UNIVERSITY OF CHICAGO

To Write or Not to Write: The Impact of Modality on Creative Problem-Solving Ability

By

Nairi Brown

June 2021

A paper submitted in partial fulfillment of the requirements for the
Master of Arts degree in the
Master of Arts Program in the Social Sciences

Faculty Advisor: Boaz Keysar
Preceptor: Samantha Fan

Abstract

Spoken and written presentations can differ on a variety of dimensions. While spoken presentation may promote intuitive thought processes, it can also increase cognitive load. Given that creative problem-solving tasks are known to be aided by intuitive thinking, varying the modality of the presentation could impact creative problem-solving ability. We designed a study to evaluate the way in which modality impacts creative problem-solving ability by altering the presentation modality of a creative problem-solving task between subjects. 192 Native English speakers completed the Remote Associate Task (RAT) in either written or spoken form, subsequently completing a surprise recall task and the Cognitive Reflection Task (CRT). Participants did not differ in accuracy on the RAT based on condition, but participants in the spoken condition performed significantly better on the recall task, while participants in the written condition performed significantly better on the CRT. Our results suggest that modality may still impact performance, but only in certain dimensions.

Problem solving is characterized by effortful and controlled thinking, which involves analytical processes that rely on working memory and focused attention (Aiello et al., 2012). For example, researchers have argued that analytical thinking is required for ambiguous situations in which learners must create or identify an unsolved problem (Robbins, 2011). In general, it is assumed that focused and controlled attention aids problem-solving ability (Aiello et al., 2012) by reducing vulnerability to heuristics and biases (Pretz, 2008). An effortful and controlled thinking process is contrasted with one that is more intuitive and passive. A more intuitive thinking process is one that is effortless, affective, slower, context specific, and automatic (Pretz, 2008). Such a process relies less on cognitive control and focused attention than does an effortful, analytical approach (Aiello et al., 2012).

Exceling at creative problem-solving tasks requires effortful and controlled thinking (Dietrich, 2004). In particular, the Remote Associate Task (RAT) has been used as a measure of creative problem-solving in many different experiments (Leung et al., 2012; Miron-Spektor, Gino, & Argote, 2011; Roskes, De Dreu, & Nijstad, 2012). In the RAT, participants are given three words (ex: cream, skate, water) and asked to determine a fourth word that connects to the three words (ex: ice). The RAT has been considered a reliable measure of creative problem-solving because information that is initially retrieved and is most related to the words is often incorrect. Instead, individuals must rely on more distantly-related information to find a connection to all three words (Bowden & Jung-Beeman, 2003). If effortful and controlled thinking enhances creative problem-solving ability, then this type of thinking process should enhance performance on the RAT.

Interestingly, recent studies have demonstrated that a less analytical and controlled solution approach may also enhance performance on creative problem-solving tasks. Aiello et al.

(2012) asked half of 153 monolingual English-speaking participants to "use [their] gut" before solving the RAT triads, while half received no instruction to rely on gut feelings. Participants who were told to rely on gut instincts outperformed those who were not told to rely on gut instincts. The authors argued that the "use your gut" instruction shifted participants away from an effortful, analytic solution approach and instead encouraged more passive, intuitive processing. Therefore, a less controlled, more intuitive thinking process may actually improve performance on creative problem-solving tasks.

At the same time, there are limits to the benefits of intuitive reasoning on creative problem-solving ability. In some situations, particularly those in which cognitive load is high, intuitive and passive reasoning seems to have no effect on creative problem-solving ability. Thus, cognitive load may overtake the effects of intuitive thinking. Along with the group of monolinguals, Aiello et al. (2012) studied a group of 141 early bilingual or non-native English speakers. The authors found that when the early bilingual and non-native English speakers completed the RAT in a non-native language, they were significantly outperformed by monolinguals in the intuition condition ("use your gut"). However, this disadvantage was not observed in bilinguals who achieved native fluency in their second language, and presumably had to work less to access English words, which may have led to a reduced cognitive load as compared to non-native speakers. Geipel et al. (2016) also suggested that the comprehension of a second language requires more cognitive resources than the comprehension of a native language. Taken together, these results indicate that the benefits of intuitive reasoning on creative problem-solving tasks (specifically, the RAT) may be differentially impacted by other factors, such as enhanced cognitive load.

Relatedly, using and interpreting spoken language places a higher cognitive load on the individual than processing written language. Pupil dilation is an established and valid measure of cognitive load, with greater pupil dilation reflecting greater cognitive load (Klingner et al., 2011). Researchers found significantly greater pupil dilation when mental multiplication, digit span memory, and vigilance tasks were presented in spoken form rather than visually (Klingner et al., 2011). These findings indicate that presenting information in spoken form increases cognitive load, which decreases performance on reasoning tasks. Though participants in Klingner et al. (2011) completed a reasoning task that relies on associations between numbers, it is reasonable to expect that the impact of modality may extend to a reasoning task that relies on associations between words (the RAT). Therefore, the way in which a problem is presented may play a role in impacting creative problem-solving ability.

However, a recent study has shown that spoken presentation could, in fact, enhance intuitive reasoning, which could in turn affect performance on creative problem-solving tasks. This result complicates the work of Klingner et al. (2011), who showed that spoken presentation only enhances cognitive load, impairing performance on creative problem-solving tasks. Geipel & Keysar (2021) gave participants different reasoning paradigms, including semantic distortion detection, insight problems, and syllogism validity judgments, in either spoken or written form. When information was presented in spoken form, participants responded more intuitively than when the same information was presented in written form. Participants also completed a Cognitive Reflection Task (CRT) which consists of questions that are inhibited by intuitive reasoning. Participants in the spoken condition were significantly more likely to give the incorrect, intuitive answer than those in the written condition. As a result, reasoning based on spoken language may rely more heavily on intuition than reasoning based on written text.

Altogether, these studies have suggested a positive correlation between intuitive reasoning and creative problem-solving ability, but this correlation may be impacted by the modality in which the problem is presented. If modality does impact intuitive reasoning, then it may also impact creative problem-solving ability.

The current study seeks to directly assess how the modality in which information is presented (written or spoken) impacts creative problem-solving ability. Native English speakers completed the RAT with the words presented either in written or spoken form. We then compared the results between the two groups. If spoken language, as opposed to written language, promotes more intuitive reasoning, and intuitive reasoning promotes creative problem-solving, then participants' performance on the RAT should increase when words are presented in spoken form. This prediction is known as the **Process Engagement Account**. However, if using and encoding spoken language is harder and more resource-depleting than encoding written language, then participants' creative problem-solving ability should be diminished when information is presented in spoken form. This prediction is known as the **Cognitive Load Account**.

Even though the information required to make important decisions may vary in the mode of presentation, it remains unclear as to whether spoken or written information enhances creative problem-solving ability. Clarifying how presentation impacts performance on the RAT will help clarify how modality impacts creative problem-solving ability.

Method

Participants

We recruited 350 participants through the Center for Decision Research at the University of Chicago. Previous studies have indicated that performance on the RAT is impaired when the

RAT is presented in a non-native language (Aiello et al., 2012). Therefore, we only recruited native English speakers. Participants could not have hearing impairments. Participants could not start the task if they answered an audio/visual test question asked at the beginning of the experiment incorrectly twice, indicated that they were a non-Native English speaker, or failed to correctly answer what the task was twice. We excluded a total of 170 participants. We planned to exclude participants who scored two standard deviations (SD) below the mean accuracy[1], participants who indicated technical difficulties (n = 27), participants who could not remember a single word in the recall task (n = 43), participants who indicated that they had prior experience with the RAT task (n = 34), participants who indicated that they did not pay attention throughout the task (n = 11), and participants who indicated that they wrote down the words during the RAT task (n = 88). Some participants were excluded for multiple reasons. These exclusions were pre-registered and were made prior to any data analyses. This resulted in a final sample of 192 participants ranging from the age of 18-60 years (M = 26.34, SD = 8.88). 70 participants (54 female, 15 male, 1 other) completed the spoken condition and 122 participants (82 female, 37 male, 3 other) completed the written condition. Participants in the spoken condition were significantly older than participants in the written condition (p = .02). However, a simple linear regression showed that the relationship between age and accuracy on the RAT triads was not significant ($F(1, 190) = 0.07$, p = .79, $\eta^2 < 0.01$).

Materials

We hired four male native English speakers who were unaware of the purpose of the experiment through the Fiverr online marketplace. Speakers had a standard American accent, and recorded 48 words from 16 word triads. Speakers were told to read each word aloud in a neutral

[1] We did not end up excluding these participants in our comparison of accuracy between conditions. Including these outliers did not change the conclusion that RAT accuracy does not vary based on condition (p = .22).

tone. We normalized the loudness and peak amplitude of each speakers' recordings using version 2.4.2 of the Audacity (R) software (Audacity Team, 2020). We also removed the silence from each word spoken to ensure that the recordings began playing exactly at the onset of the word, and stopped playing exactly at the offset of the word. We used these recordings for the spoken condition. For the written condition, the words were presented in black in the center of a white screen. We matched the onset and duration of the words for the spoken and written conditions, using the methods outlined in Klingner et al. (2011).

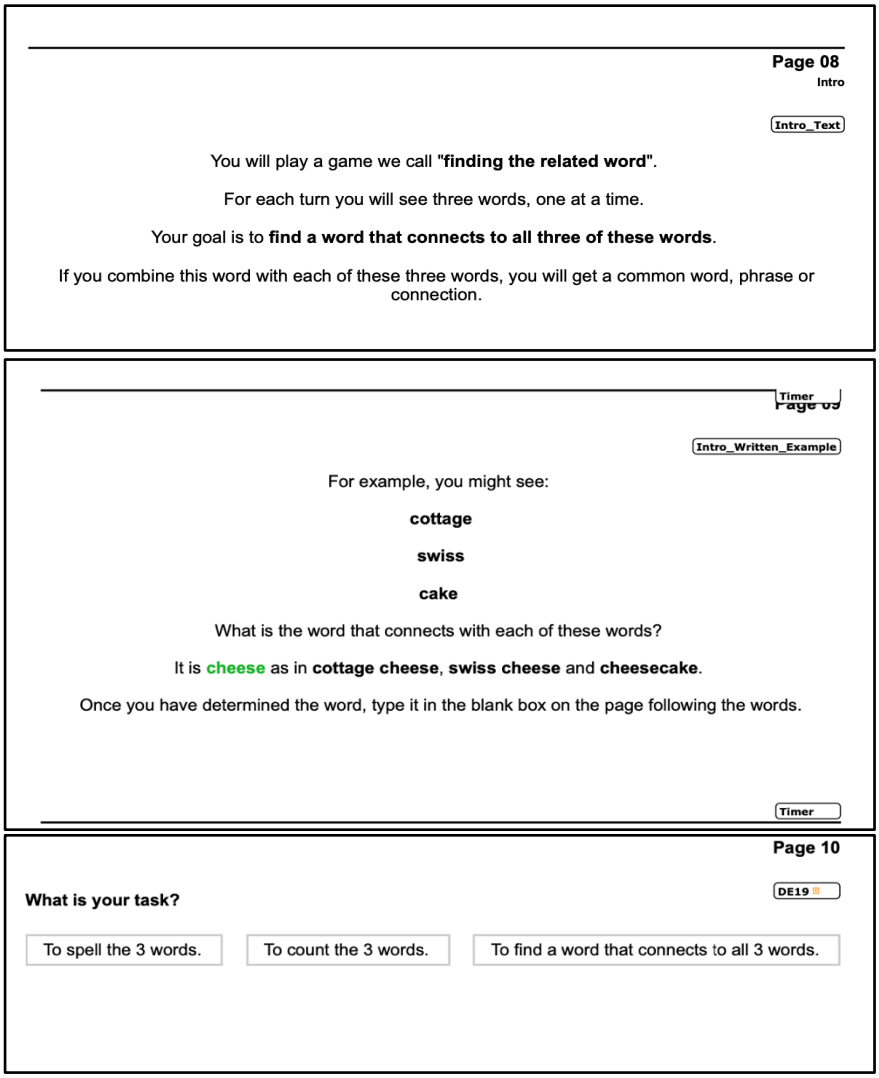
Procedure

Participants were presented with 16 word triads that were displayed in a random order (Table 1). Before beginning the actual test, participants were told that their goal was to find a word that connected to all three words presented. They were told that if they combined the connecting word with each of the three words shown, they would get a common word, phrase, or connection (Figure 1). Participants were randomly assigned to either the spoken or written condition.

Participants were then given an example of a RAT triad (cottage, swiss, cake) in either spoken or written form (depending on assigned condition). They were then explicitly shown the correct answer to this question (cheese). Participants were also asked to indicate what they understood to be the task (Figure 1). As discussed, participants who failed to give the correct answer twice were not allowed to start the task. Participants then completed two trial RAT triads that did not count in the final analysis to orient participants to the task. In these trials, the connecting word was eventually revealed (Figure 1).

Figure 1

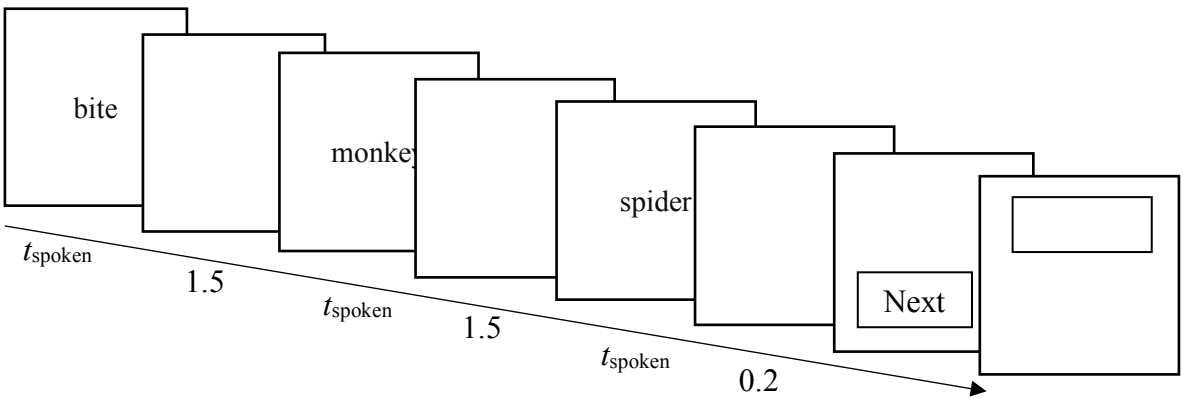
Training Phase



In both conditions, the words were separated by a blank white screen that lasted for 1.5 seconds. After presentation of the third word, a “next” button appeared on the screen, taking participants to a blank box. Participants were asked to type the word that connected the three prior words into the blank box (Figure 2).

Figure 2

Experimental Design



The last triad presented (sandwich, golf, foot) was the same for all participants in both conditions. After determining the connecting word for this triad, participants were unexpectedly asked to recall the three words of this final triad. This test was used to determine if one modality differentially impacts recall ability. We then asked participants how easily the connecting words came to mind. Participants answered using a 1-6 scale, with 1 representing "not at all easily", and 6 representing "very easily". We also asked participants if they had any experience with the RAT task prior to the study.

Participants then completed the CRT in which the questions were presented in written form. The CRT consists of questions that are inhibited by intuitive reasoning. We used responses to the CRT to account for differences in individual performance, assessing if performance on the CRT predicts performance on the RAT. We also used these responses to assess if the modality of presentation of the RAT impacts performance in a reasoning task that follows.

Results

We tested the impact of language modality on performance in the RAT triad task. We averaged performance on each of the 16 RAT triads for each participant. Each participant received a 1 for determining the correct connecting word and a 0 for giving an alternative answer. This led to an average score between 0-1 for each participant in each condition, with

higher scores reflecting greater performance on the RAT task. We labeled this variable as "Accuracy".

Mean accuracy across all participants on all trials was relatively low, with participants answering less than half of the triads correctly on average ($M = 0.41$, $SD = 0.24$). Mean accuracy across all participants on all trials by condition was also relatively low. Participants in the written condition answered less than half of the triads correctly ($M = 0.41$, $SD = 0.23$), and participants in the spoken condition performed similarly ($M = 0.41$, $SD = 0.26$). Participants did not differ in accuracy based on condition ($p = .60$). A summary of the mean accuracy on each trial by condition is provided below in Table 1.

Table 1

Average Performance on Triad by Condition

Correct Word	Written Accuracy	Spoken Accuracy
Deep	0.25	0.21
Stamp	0.11	0.16
Blast	0.18	0.23
Black	0.26	0.30
Star	0.30	0.36
Water	0.50	0.41
Cheese	0.36	0.36
Spider	0.43	0.47
Bar	0.34	0.37
Club	0.48	0.39
Hard	0.48	0.47
Sea	0.57	0.60
White	0.40	0.43
Beer	0.57	0.49
Glass	0.67	0.67
Cookie	0.59	0.61

We also analyzed average performance on the RAT triads based on the difficulty of the triads. Using average performance from a pilot study in which all participants received the RAT

triads in written form, we classified the 8 triads with mean accuracy above 50% as easy, and the 8 triads with mean accuracy below 50% as hard. The classifications of triads as easy or hard, along with the average accuracy, are provided below in Table 2. We confirmed that these classifications aligned with the current data using a Wilcoxin signed-rank test, which revealed that accuracy was significantly higher for easy word triads than for difficult word triads (M_{easy} = 0.30, $M_{hard.}$ = 0.51, $p < .001$).

Table 2

Classifications of Triad Difficulty

Correct Word	Mean Accuracy	Classification
Deep	0.32	Hard
Stamp	0.33	Hard
Blast	0.35	Hard
Black	0.38	Hard
Star	0.38	Hard
Water	0.43	Hard
Cheese	0.47	Hard
Spider	0.50	Hard
Bar	0.53	Easy
Club	0.53	Easy
Hard	0.57	Easy
Sea	0.57	Easy
White	0.58	Easy
Beer	0.62	Easy
Glass	0.62	Easy
Cookie	0.63	Easy

Using a Mann-Whitney U test, we determined that average performance on the easy RAT triads was not significantly impacted by modality ($M_{written}$ = 0.51, M_{spoken} = 0.51, $SD_{written}$ = 0.25, SD_{spoken} = 0.28, W = 4300, p = .74). We also determined that that average performance on the hard RAT triads was not significantly impacted by modality ($M_{written}$ = 0.30, M_{spoken} = 0.31, $SD_{written}$ = 0.26, SD_{spoken} = 0.21, W = 4391.5, p = .95). To confirm this result, we ran a binary

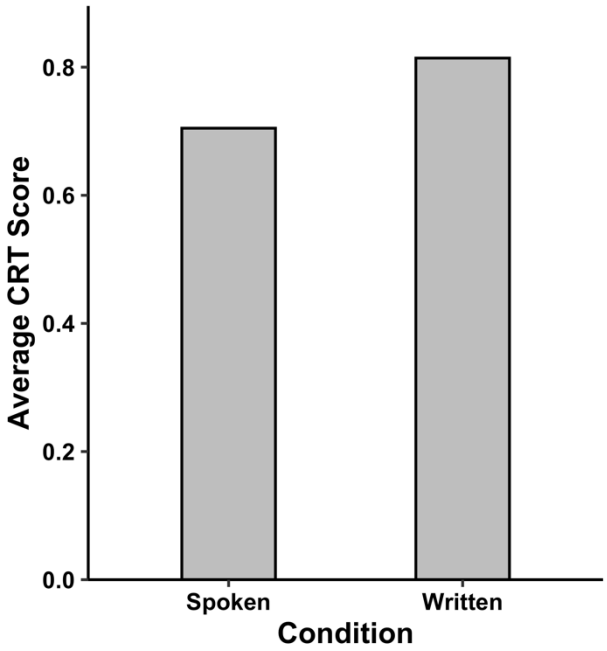
logistic regression, treating accuracy across triads as the outcome variable (0 = incorrect, 1 = correct) and difficulty (0 = difficult, 1 = easy), modality (0 = written, 1 = spoken), and the difficulty by modality interaction term as predictors. The regression indicated that there was no significant interaction between modality and difficulty ($z = 0.73$, $p = .47$, $\beta = .11$, $OR = 1.12$). The regression also revealed that modality and difficulty together predicted accuracy ($R^2_N = .06$, $p < .001$), yet individually, only difficulty significantly predicted accuracy ($z = 6.46$, $p < .001$, $\beta = .80$, $OR = 2.23$). By contrast, modality did not significantly predict accuracy ($z = –0.48$, $p = .628$, $\beta = –.05$, $OR = 0.95$).

Performance in the Recall Task

We then analyzed the impact of modality on performance in the surprise recall task given at the end of the 16 RAT triads (ability to recall the three words of the last triad). Using a Mann-Whitney U test, we determined the difference in performance between conditions. Specifically, participants in the spoken condition correctly recalled significantly more words of the last triad ($M = 2.83$, $SD = 0.45$) than participants in the written condition ($M = 2.57$, $SD = 0.59$, $W = 3286$, $p < .001$). This result suggests that when the RAT task is presented in spoken form, subsequent recall rates may be improved as opposed to when the RAT is presented in a written form.

Figure 3

Recall Ability by Condition



Ease of Responding

We then determined whether participants differed in self-report of how easily the words came to mind based on condition. Using a Mann-Whitney U test, we determined that the difference in reported ease between conditions was not significant ($M_{written}$ = 2.48, M_{spoken} = 2.31, $SD_{written}$ = 1.18, SD_{spoken} = 1.04, W = 4545, p = .44).

Performance in the CRT Task

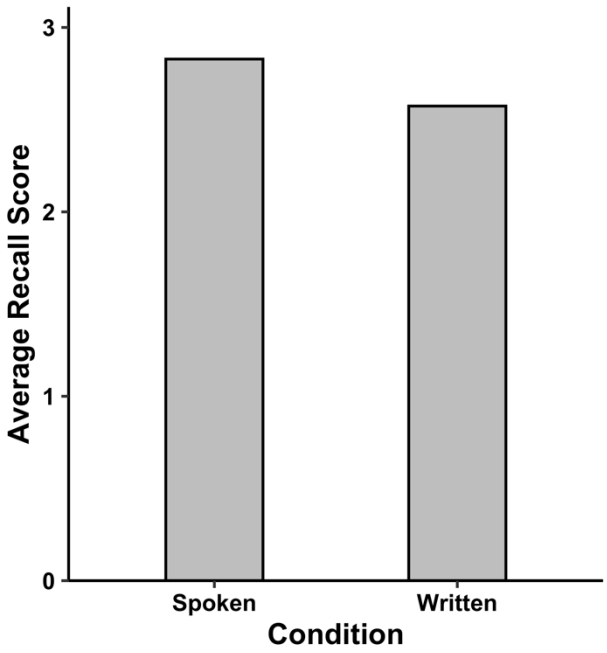
The CRT consisted of three reasoning questions. Each participant in each condition received a 1 for determining the correct answer and a 0 for giving an incorrect answer. We treated accuracy as the mean score across the three questions, which resulted in a score between 0-1. We labeled this variable as "Average CRT".

Mean accuracy across all participants was relatively high (M = 0.77, SD = 0.29). Participants in the written condition had a greater average performance on the CRT (M = 0.81, SD = 0.26) than participants in the spoken condition (M = 0.70, SD = 0.32). The Mann-Whitney

U test for the relationship between condition and performance on the CRT indicated that the difference in performance between conditions was significant (p = .02). This result suggests that spoken presentation may have promoted a more intuitive thought process. A more intuitive thinking process would hinder performance on the reasoning tasks, because these tasks require the participant to override an initial, intuitive answer. Therefore, the reduced performance in the spoken condition suggests that participants in this condition may have been engaging in more intuitive thinking.

Figure 4

CRT Performance by Condition



Given that participants in the spoken condition were significantly older than those in the written condition, we assessed whether CRT performance was predicted by modality, age, and the interaction between modality and age. A multiple linear regression model showed that the

interaction between modality and age significantly predicted accuracy on the CRT ($\beta = -.71$, $F(1, 187) = 8.45$, $p = .004$, $R^2\Delta = .04$). Age alone did not significantly predict accuracy on the CRT ($F(1, 190) = 0.07$, $p = .79$, $\eta^2 < 0.01$)). Using the Johnson-Neyman method and 10,000 bootstrapped samples, we determined that the effect of modality on performance was statistically significant for participants older than 25 years (38.7%). Studies have shown that adults tend to show a cognitive decline with age. Though the actual onset of this decline is debated, some have even argued that certain aspects of decline begin in one's 20s and 30s (Salthouse, 2009). Regardless of exact onset, the correlation between aging and cognitive decline suggests that decreased performance in the spoken condition could have resulted from the significantly higher age of participants in the group.

Discussion

The current study sought to analyze the way in which presentation modality (spoken or written) impacts performance on a creative-thinking task. We also addressed the way in which modality may impact other relevant processes – namely recall ability and performance on reasoning tasks.

Overall, our results did not indicate a significant difference in performance on the RAT triads between participants who were in the spoken condition and participants in the written condition. As a result, we do not have conclusive evidence to support either the Process Engagement Account or the Cognitive Load Account at this time. Interestingly, we did find significantly enhanced performance for participants in the spoken condition on the surprise recall task. We also found that participants in the written condition outperformed those in the spoken condition on a subsequent reasoning task presented in written form for all participants.

Null findings are often challenging to interpret. However, we provide a few speculations and future directions. First, it is worth noting that this study could have suffered from a lack of power. A theoretical power analysis revealed a requirement of at least 274 participants. Initially, we sought to recruit 300 participants, which would have allowed for up to 9% data loss, or data loss from 26 participants. However, due to our strict preregistration inclusion criteria, we were only able to use data from 192 participants, significantly reducing the power of the study. Had we maintained a higher power level, it is possible that we could have detected differences between participants in the spoken and written conditions.

Second, it is worth noting that performance on the creative-thinking task was generally low in both conditions, with all participants showing an average accuracy below 0.5. As a result, it is possible that the RAT task was too cognitively demanding to detect differences between participants in the spoken and written conditions. That is to say, it is possible that the RAT triads induced such a high cognitive load on all participants that differences due to intuitive reasoning were overpowered or the additional cognitive load generated by the spoken condition could not compare to the cognitive load of the task in general. The findings of Aiello et al. (2012) support this claim, as the researchers found that participants who completed the RAT in a non-native language, which is more cognitively demanding than completing the task in a native language, were significantly outperformed by monolinguals in the intuition condition. Therefore, the impact of intuition was overpowered by the additional cognitive load. In the current study, it is possible that the cognitive load of the RAT task – analogous to the cognitive load of processing a non-native language – masked the impact of intuitive reasoning.

In our exploratory analyses, we did not find differences in performance based on the difficulty of the RAT triad. Given that the RAT triads appeared to be difficult overall, this

finding is not surprising. It is likely that RAT triads labeled as "easy" were only easy relative to the harder triads, but were still highly cognitively demanding for participants. As a result, we would not expect to see differences when controlling for difficulty.

Assuming our null findings to be true, our results challenge both the Process Engagement Account and the Cognitive Load Account. In contrast to these theories, our results suggest that the modality in which a creative problem-solving task is presented may not influence overall performance. On the other hand, it is also possible that *both* the Process Engagement Account and the Cognitive Load Account may account for separate aspects of the creative problem-solving process. Should this be true, it is possible that spoken language does in fact enhance intuitive reasoning while also increasing cognitive load, but the benefit of greater intuitive reasoning was overridden by the detriment induced by increased cognitive load. Such an interpretation and speculation would align with our finding that participants in the written condition outperformed those in the spoken condition on the CRT, a task that is impaired by intuitive reasoning.

Though we did not find differences in performance on RAT triads, participants in the spoken and written conditions did differ on some dimensions. First, participants in the spoken condition recalled significantly more words of the final RAT triad than participants in the written condition, suggesting that the modality of presentation may in fact impact recall ability. Previous research has indicated that spoken language may enhance recall ability. For example, Conway and Gathercole (1987) found that participants who either heard or spoke words aloud showed significantly enhanced memory than participants who read the words silently on a delayed surprise memory test. Interestingly, our finding also contrasts with some recent literature on the impact of modality on memory. For example, Kessels (2003) found that patients who received

medical advice in written form retained that information better and adhered to treatment plans more often than patients who received medical advice in spoken form. Though this specific study may not map on well to the design of our study, it furthers the idea that modality may differentially impact recall. Future studies could evaluate the way in which modality may impact recall ability, especially for highly consequential information, such as medical treatment guidelines.

Participants also differed in performance on a reasoning task that followed the creative problem-solving task. While participants in the spoken condition outperformed those in the written condition on surprise recall, participants in the written condition outperformed participants in the spoken condition on the CRT task. Recall ability and performance on the CRT were not strongly correlated. As discussed, it has been found that spoken language promotes a more intuitive thinking process (Geipel & Keysar, 2021). Given that success on the CRT requires one to overcome an initial, intuitive answer, the tendency of spoken language to promote intuitive reasoning would inhibit performance on this task. Therefore, our results may align with the notion that spoken language may promote a more intuitive thinking process, though we cannot confidently make such a claim. Regardless of mechanism, these results show that the modality in which a creative problem-solving task is presented may impact performance on subsequent reasoning tasks.

We also found that the interaction between modality and age significantly predicted accuracy on the CRT. Specifically, the effect of modality on performance was significant for individuals over 25 years of age, and accounted for 38.7% of the variability in performance. The cognitive decline associated with age may have contributed to decreased performance. As discussed, many studies have found that adults show a cognitive decline with age, with certain

aspects of decline beginning as early as in one's 20s and 30s (Salthouse, 2009). Given that individuals in the spoken condition were significantly older than individuals in the written condition, this finding aligns with the result that individuals in the written condition outperformed those in the spoken condition on the CRT. In the future, studies should replicate the RAT and CRT tasks with participants in a smaller age range to reduce the impact of age on performance. Doing so would help elucidate whether spoken language promoted intuitive reasoning, leading to reduced performance on the CRT, or if age primarily accounted for the difference in performance. Given that age did not significantly predict performance on the RAT task itself, we argue that it is unlikely that age was the primary contributor to reduced performance in the spoken condition. Altogether, the present results do not provide sufficient evidence for either the Process Engagement Account or the Cognitive Load Account. However, we did find significant differences on various dimensions between individuals in the written and spoken conditions. Specifically, modality appears to impact both recall ability and performance on reasoning tasks that follow a creative thinking task. Additionally, age interacts with modality to impact performance on the reasoning task.

The way in which individuals receive information constantly varies. Our results suggest that individuals may want to place focus not only on the modality in which the information is presented, but on the required tasks that follow the presentation. For example, if retention is a key consideration, individuals may want to consider presenting the information in spoken form. By contrast, if subsequent reasoning ability is paramount, individuals may want to consider presenting the information in written form. Though our findings did not provide conclusive evidence for the Process Engagement Account or the Cognitive Load Account, we maintain that modality should, in theory, impact subsequent reasoning ability. Future studies should clarify the

way in which modality specifically impacts reasoning ability so as to better understand when one modality may prove advantageous.

References

Aiello, D. A., Jarosz, A. F., Cushen, P. J., & Wiley, J. (2012). Firing the executive: When an analytic approach to problem solving helps and hurts. *The Journal of Problem Solving, 4*(2), 7.

Audacity Team. (2020). *Audacity Team* (p. Last viewed September 20, 2014). https://audacityteam.org

Bowden, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. *Behavior Research Method, Instruments, & Computers*, *35*(4), 634–639.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *26*(3), 341-361.

Dietrich, A. (2004). The cognitive neuroscience of creativity. *Psychonomic bulletin & review*, *11*(6), 1011-1026.

Geipel, J., Hadjichristidis, C., & Surian, L. (2016). Foreign language affects the contribution of intentions and outcomes to moral judgment. *Cognition*, *154*, 34-39.

Geipel, J., & Keysar, B. (2021). *Listening Speaks to our Intuition while Reading Promotes Analytic Thought.* Manuscript submitted for publication.

Kessels, R. P. (2003). Patients' memory for medical information. *Journal of the Royal Society of Medicine*, *96*(5), 219-222.

Klingner, J., Tversky, B., & Hanrahan, P. (2010). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), 323–332. w

Leung, A. K. Y., Kim, S., Polman, E., Ong, L. S., Qiu, L., Goncalo, J. A., & Sanchez-Burks, J. (2012). Embodied metaphors and creative "acts". *Psychological science, 23*(5), 502-509.

Mednick, S. (1962). The associative basis of the creative process. *Psychological review, 69*(3), 220.

Miron-Spektor, E., Gino, F., & Argote, L. (2011). Paradoxical frames and creative sparks: Enhancing individual creativity through conflict and integration. *Organizational Behavior and Human Decision Processes, 116*(2), 229-240.

Robbins, J. K. (2011). Problem solving, reasoning, and analytical thinking in a classroom environment. *The Behavior Analyst Today*, *12*(1), 41.

Roskes, M., De Dreu, C. K., & Nijstad, B. A. (2012). Necessity is the mother of invention: Avoidance motivation stimulates creativity through cognitive effort. *Journal of personality and social psychology, 103*(2), 242.

Pretz, J. E. (2008). Intuition versus analysis: Strategy and experience in complex everyday problem solving. Memory & cognition, 36(3), 554-566.

Salthouse, T. A. (2009). When does age-related cognitive decline begin?. *Neurobiology of aging*, *30*(4), 507-514.

Appendix

Cognitive Reflection Task (CRT) Questions

"Ants are walking in a line. A rude ant cuts in front of the ant walking second in line. What is the rude ant's place in the line now?"

"A farmer had 15 cows and all but 8 died. How many are left?"

"John and Peter each have 10 ties. John gives Peter 2 ties. How many more ties does Peter have than John?"