Sight Over Sound: Working Memory Differences Between Modalities in Decision-Making and Reasoning

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Arts
Department of Psychology: Cognition and Neuroscience Program
University of Saskatchewan
Saskatoon

By

Issac Pasloski

© Copyright Issac Pasloski, November 14th, 2022. All Rights Reserved. Unless otherwise noted, the copyright of the material in this thesis belongs to the author.
Permission To Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate Master of Arts from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying, publication, or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use that may be made of any material in my thesis.

Disclaimer

Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Psychology Department
154 Arts, 9 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5A5, Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan, S7N 5C9, Canada
Abstract

Every day we utilize the stimuli around us to make decisions, mainly from the visual and auditory modalities. However, research methods investigating the cognitive processes that are involved in decision-making and reasoning processes only utilize a visual presentation of stimuli. Research shows that there is a distinct difference between the encoding and processing of stimuli between the auditory and visual modalities. This difference in stimulus processing may impact the effectiveness of the working memory storage that is used in decision-making and reasoning.

This thesis utilized two separate experiments comparing the performance of the auditory and visual modalities on the base rate neglect task. Experiment 1 was not sensitive enough to any working memory differences between modalities. Experiment 2, using a secondary memory load task, did detect differences in working memory capacity, but only within the secondary task. This research shows that there is a difference in the working memory load capacity between modalities, which can be addressed in future studies with more rigorous measurements of cognitive load and provide more insight into modality differences in decision-making and reasoning.

**Keywords:** Base rate neglect, cognitive load, auditory modality, visual modality, working memory, decision-making, reasoning, memory
Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Steven Prime, for everything he has done over these last few years. Despite the immense trials and difficulties faced during this time he has adapted just as much as I have, and acted as a guide to ensure the completion of this thesis. Without his encouragement, intelligence, and aid in the cultivation of curiosity in myself, I could not have completed this project without him.

I also would like to thank the members of my advisory committee, Dr. Valerie Thompson, and Dr. Janeen Loehr, for their advice and immense store of knowledge which helped me push the boundaries of what I know in research. Thanks also to Dr. Marla Mickleborough for being my external in my defence and helping me reflect on my journey. I also thank Dr. Ron Borowsky for their early input and critique on this project.

I would also like to thank T-Bone from Sound Lounge by T-Bone for being the voice actor in the auditory stimuli for this project. His expertise in sound design and voice acting made the more daunting parts of this project easy.

Finally, I would like to thank my family, specifically my parents, Michael Pasloski and Brenda Pasloski, as well as my grandma, Anne Pasloski. All of whom since childhood have aided and supported me in my journey of science and academia; helping my curiosity grow to new higher heights and always believing in me.
## TABLE OF CONTENTS

PERMISSION TO USE................................................................................................................. i
DISCLAIMER................................................................................................................................ i
ABSTRACT................................................................................................................................... ii
ACKNOWLEDGEMENTS ........................................................................................................ iii
TABLE OF CONTENTS ........................................................................................................ iv
LIST OF FIGURES ................................................................................................................... vii
CHAPTER 1 .................................................................................................................................. 1
  General Introduction .............................................................................................................. 2
    Processing in Decision-Making and Reasoning ................................................................. 3
      Base Rate Neglect ........................................................................................................... 4
  Working Memory................................................................................................................... 6
    Foundational Working Memory Literature ...................................................................... 6
    Working Memory Subcomponents ..................................................................................... 7
DIFFERENCES IN MODALITY PERFORMANCE ....................................................................... 8
  Auditory Performance Superiority ..................................................................................... 9
  Visual Performance Superiority ........................................................................................... 9
  Serial Position Effect .......................................................................................................... 10
Present Research .................................................................................................................. 11
CHAPTER 2 ................................................................................................................................ 12
Experiment 1 ........................................................................................................................... 13
Hypotheses .............................................................................................................................. 13
  Hypothesis A ....................................................................................................................... 13
  Hypothesis B ....................................................................................................................... 14
  Hypothesis C ....................................................................................................................... 14
  Hypothesis D ....................................................................................................................... 14
Methods .................................................................................................................................... 15
  Subjects ............................................................................................................................... 15
  Apparatus ........................................................................................................................... 15
  Stimuli ................................................................................................................................... 16
  Procedure ............................................................................................................................ 16
Results ....................................................................................................................................... 21
  Base Rate Neglect Accuracy ............................................................................................... 21
    Congruency ....................................................................................................................... 22
    Base Rate Position ........................................................................................................... 22
    Modality .......................................................................................................................... 22
    Interaction Effects ........................................................................................................... 22
  Reaction Times .................................................................................................................. 24
CHAPTER 4

General Discussion ................................................................. 51
Experiment 1 Findings ........................................................... 52
  Base Rate Neglect Performance ........................................... 52
Experiment 2 Findings ........................................................... 53
  Base Rate Neglect Performance ........................................... 54
  Memory Task Performance .................................................. 54
Between Experiment Results .................................................. 56
Limitations and Future Directions ........................................... 57
Conclusion ............................................................................... 60
References .............................................................................. 61
Appendices .............................................................................. 67
  Appendix A .......................................................................... 67
  Appendix B .......................................................................... 72
  Appendix C .......................................................................... 73
  Appendix D .......................................................................... 80
### List of Figures

- Figure 1. General experimental paradigm for the visual group in Experiment 1 ......................... 19
- Figure 2. General experimental paradigm for the auditory group in Experiment 1 ..................... 20
- Figure 3. Mean probability estimates for Experiment 1 as a function of Base Rate Ratio Position x Base Rate Problem Congruency ................................................................. 23
- Figure 4. General experimental paradigm for the visual group in Experiment 2 ......................... 33
- Figure 5. General experimental paradigm for the auditory group in Experiment 2 ..................... 34
- Figure 6. Mean probability estimates for Experiment 2 as a function of Memory Load size x Base Rate Problem Congruency. ........................................................................... 37
- Figure 7. Mean memory load accuracy for Experiment 2 as a function of Array Size x Modality ................................................................................................................................... 40
- Figure 8. Mean memory load accuracy for Experiment 2 as a function of Base Rate Problem Congruency x Modality .......................................................... 41
- Figure 9. Cowan’s K value for Experiment 2 as a function of Base Rate Problem Congruency x Modality x Array Size................................................................................................ 44
- Figure 10. Reaction time for the memory load task for Experiment 2 as a function of Base Rate Problem Congruency x Modality x Array Size............................................. 46
Chapter 1:

Introduction
General Introduction

A variety of approaches in research have investigated the way people reason or make decisions. Despite the abundance of methodologies, the overwhelming majority of tasks have been in the visual modality; in the form of word-based problems. This unimodal approach limits the extent to which we can fully understand decision-making and reasoning, since in the real world decision-making problems are also encountered orally and there are distinct differences between auditory and visual processing. It is known that information processing differs on the basis of presentation modality (Paivio, 1991). Differences in stimulus processing exhibit differences in performance on a variety of cognitive tasks such as recognition, recall, arithmetic, and comprehension (Smith & Fogerty, 2015; Amon & Bertenthal, 2018; Buchweitz et al., 2009; DeStefano & LeFevre, 2004; Klingner et al., 2011). This field of research is divided in establishing which modality yields general superior performance across cognitive tasks, as each modality has been shown to lead to better performance depending on the task. This previous research has been limited to studies in simple cognitive tasks. It remains unclear how stimulus presentation may affect reasoning. Therefore, the purpose of this research is to understand what impact the presentation modality of a cognitive reasoning task will have on decision-making and reasoning.

The use of a unimodal approach to study decision-making and reasoning is incorrect without proper investigation of the effect presentation modality may have. To properly explore this possibility, an overview of relevant decision-making and reasoning literature focused on the dual process theory and the base-rate neglect task will be discussed. Following this, a review of working memory literature to understand the cognitive processes involved in the perception, processing, and manipulation of information based on the presentation modality of stimuli.
Processing in Decision-Making and Reasoning

The dominant theory in understanding the processes behind decision-making and reasoning, how information is manipulated and processed, resulting in the outcome of a decision or answer to said problem, is known as the dual process theory. Dual process theory divides the cognitive processing and response in decision-making and reasoning into two processes, Type 1 and Type 2, each of which has their own set of defining features (Evans & Stanovich, 2013; Kahneman, 2003; Sloman 1996). Type 1 processing is characterized as autonomous activation and minimal usage of working memory (Evans & Stanovich, 2013). This would be characterized as saying the first thing that comes to one’s mind when asked a question; an initial “gut” response. Type 2 processing involves a greater usage of working memory due to the cognitive load involved in cognitive decoupling (Evans & Stanovich, 2013). This cognitive decoupling is the ability to divide a larger piece of information into its subcomponents and evaluate each smaller fragment as an individual item.

A large portion of the research investigating decision-making and reasoning from the framework of dual process theory use test items or problems that, at first glance, appear to have a quick response, either due to the formatting of the question, the accessibility of relevant information for making a decision, or the contextual information, resulting in fallible reasoning (Frederick, 2005; Kahneman, 2003; Toplak et al., 2011). The typical “easy” answers would be indicative of that from Type 1 processing. However, the correct answer typically requires the engagement of working memory; the engagement of Type 2 thinking to correctly process the information (Newman et al., 2017; Kahneman, 2003).
**Base Rate Neglect**

One of the well-established tasks for studying dual process theory is the base rate neglect task. Base rate neglect is a cognitive bias that occurs where people will disregard, or neglect, statistical information in the presence of personality description (Kahneman & Tversky, 1973). This is investigated using the base rate neglect task, where people are given a base rate of group populations within a sample, and a description of a random individual of that sample and then asked to estimate the probability the described individual is from a certain population (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). A typical problem presented to subjects in a base rate neglect task is shown below:

In a study of 1000 people there were 3 engineers and 997 lawyers. Jack is a randomly selected participant of this study. He spends most of his time on his hobbies such as carpentry, sailing, and mathematical puzzles.

What is the probability that Jack is an engineer?

Rather than use the base rate, subjects tend to base their probability judgements of group membership on the personality descriptions (Kahneman & Tversky, 1973; Tversky & Kahneman, 1974; Barbey & Sloman, 2007). Base rate neglect has been postulated to reflect subjects relying on heuristic Type 1 processing due to the biasing personality description in these problems (De Neys, 2006). A lack of base rate neglect is typically taken as evidence that analytical Type 2 processing has been utilized, suppressing the heuristic response (Stanovich & West, 2000). Stanovich and West (2000) posit that the engagement of Type 2 processing relies on having access to ample working memory resources, whereas without suitable cognitive load capacity, one is more likely to rely on a heuristic Type 1 based response.
The congruency of base rate problems affects the probability of conflict detection. Base rate neglect typically occurs in incongruent descriptions, where the personality description is stereotypical of the smaller group population. By changing the congruency of the base-rate problem it can act as a control. Flipping the base rates for each group population (e.g., 3 engineers and 997 lawyers, to 997 engineers and 3 lawyers) creates a congruent base rate problem; the base rates coincide with the personality description, therefore removing the cognitive conflict that occurs in typical base rate neglect problems.

The usage of a set of extreme base rate ratios (997:3, 996:4, 995:5) is often done in some studies to increase the likelihood of conflict detection in reasoning (Pennycook et al., 2012). This allows researchers to separate instances of conflicted reasoning (in incongruent problems) from the non-conflict trials. Using a more moderate base rate extremity, such as 700:300, would still create conflict in reasoning for congruent trials, but to a lesser degree than extreme base rate ratios. In the present study, extreme base rate ratios were used, akin to previous base rate neglect tasks.

Another key manipulation in the base rate neglect task is the presentation order of the base rate and the personality description. Previous studies have shown that in problems where the base rate information was presented after the personality description, subjects were more likely to give more accurate probability estimations. The increase in accuracy is due to the saliency of the base rate information which increases the likelihood of conflict detection, and accuracy of probability estimation (Pennycook et al., 2015; Krosnick et al., 1990). Consistent with previous research the present studies will also manipulate presentation order where the base rate is presented either before or after the personality description. The ordering of problems was manipulated to allow for the detection of any memory-related primacy/recency effects that may
occur between modalities. The following sections will describe working memory, primacy/recency effects, and provide context on how they might impact base rate neglect performance in different modality presentations.

**Working Memory**

As described earlier, there are two stages of cognitive processing involved in the decision-making and reasoning process. There is how the information is manipulated and processed, as covered in the dual process theory, but also how information is encoded and processed. This latter stage will be discussed in the following section.

**Foundational Working Memory Literature**

Perhaps the best known and well-established model of memory that elucidates how stimulus information is stored and manipulated for use during a cognitive task is the working memory model developed by Baddeley and Hitch (1974). Their working memory model expands on the short-term memory component of the older multistore model of memory by Atkinson and Shiffrin (1968). Put simply, the multistore model divides memory into different storage buffers that information flows from the sensory store to short-term memory if attended, and then to long-term memory if rehearsed (Atkinson & Shiffrin, 1968). Short-term memory is a limited capacity and limited duration memory store for retaining information only temporarily, often for the purpose of performing a task. Baddeley and Hitch’s (1974) model proposes that the short-term memory store itself can be further subdivided into four subsystems: the phonological loop, the visuospatial sketchpad, the central executive, and later on the episodic buffer storage (Baddeley, 2011). Verbal and visuospatial stimulus information is stored in the phonological loop and visuospatial sketchpad, respectively, and their cognitive manipulation and maintenance are governed by the allocation of attentional resources to these slave systems in the central executive.
Another key distinction between the working memory model and short-term memory as described in the multistore model is that the working memory model also emphasizes that the purpose of working memory is not just to temporarily retain information, but this is also where information is manipulated, or “worked on”, during cognitive tasks. Here, I will briefly describe each component of the multistore model of working memory.

**Working Memory Subcomponents**

The phonological loop subsystem is dedicated to storage and processing of auditory and verbal information. Expanding on Baddeley and Hitch’s (1974) initial conception of working memory, Baddeley (2003) subdivides the phonological loop into two-stage processing subcomponents: 1) the precategorical acoustic store for initial, temporary storage of auditory stimuli, and 2) the subvocal rehearsal system for maintaining information from the precategorical acoustic store in working memory for recall. Mainly only attended information from the precategorical acoustic store is processed by a subvocal rehearsal system. This subvocalization can also occur with written word stimuli, resulting in reinforcement for encoding visual stimuli (Baddeley, 2003).

The visuospatial sketchpad governs the storage and manipulation of visual and spatial stimuli. Here, a stimulus’ visual features are bound to a spatial map. From a macro perspective, this system involves the development of a spatial map of the general surroundings or pertinent physical locations (Baddeley, 2012). This can also be done on a micro-scale, keeping visual information such as matrices or pictures in working memory to be recalled later on (Baddeley, 2012).

In later developments of the working memory model, the episodic buffer storage component was proposed as a “bridge” between slave systems and long-term memory. The
episodic buffer storage combines information from the phonological loop and visuospatial sketchpad into a single “episodic representation” (Baddeley, 2000). This process acts as the integration module for updating long-term memory representations with working memory information and vice versa.

The central executive has several functions in working memory: the division of attentional resources when in a dual task scenario, specifically switching the focus of processing to either the visuospatial sketchpad or phonological loop; the selective attention processing in working memory, acting as a filter to disregard noise stimuli and only attend to information of importance (Baddeley, 1996). The central executive also acts as an initiator for the episodic buffer, either for the retrieval of information in long-term memory, or to update encoded representations based on information in working memory (Baddeley, 1996).

In summary, according to the working memory model visual and auditory information are stored in different memory buffers for retaining and manipulating information during the performance of cognitive tasks. It follows that these different modality-specific memory buffers govern reasoning task performance presented in their respective stimulus modalities. However, it remains unclear whether performance in reasoning tasks, such as the base rate neglect task depends on presentation modality.

**Differences in Modality Performance**

To my knowledge, no research has investigated whether there are any differences in reasoning performance between stimulus modalities using tasks that engage dual processing. There has been an abundance of literature comparing the processing abilities of the auditory and visual modalities in more simple working memory tasks. Literature posits the difference between
modalities is task specific. The research literature in the following subsections showcase the differing performance between modalities dependent on the nature of the task.

**Auditory Performance Superiority**

There is little research that shows cognitive advantages for the auditory modality. However, the areas it does excel in are pertinent to the base-rate neglect task. Smith and Fogerty (2015) showed that sentence comprehension and recognition were better when presented in an auditory modality than visually. Conway and Christiansen (2003) also found higher accuracy on sequence recognition tasks for the auditory modality of presentation compared to tactile and visual presentations.

**Visual Performance Superiority**

The visual presentation of information has shown advantages in different tasks as well. Klingner et al. (2011) examined the differences in performance in several numerical-based tasks when presented auditorily or visually. They found visual modality superiority with higher accuracy in mental arithmetic and single digit recall compared to auditory modality presentations (Klinger et al., 2011). These findings are consistent with studies using a different arithmetic task. When presented visually with a series of subsequent additions (e.g., 5+12+8), higher accuracy occurred compared to auditory presentations (Logie et al., 1994; DeStefano & LeFevre, 2004). This advantage for visual presentation even occurs in other languages, as shown in a comparison of English speaking and Chinese speaking subjects solving single digit multiplication problems (Lefevre et al., 2001). Jaeggi et al. (2010) using the N-back task, during which subjects respond when the current stimuli matches one presented “N” items before, showed that under high task loads there was an increase in response times and reduction in accuracy in the auditory modality compared to the visual modality. Amon and Bertenthal (2018) replicated these findings, where
the auditory modality suffered in both accuracy and response time as the cognitive load (“N”) increased.

**Serial Position Effect**

Modality performance advantages do not occur only between different tasks, but can occur within a task itself. Specifically, several studies have shown modality effects in serial recall tasks (Crowder & Morton, 1969; Beaman, 2002). In a typical serial recall task, subjects are asked to memorize a list of items and followed by a test where they recall as many items as they can. The classic result in serial recall tasks is the serial position effect. The serial position effect is that item recall is highest for items presented at the beginning and the end of the list, and lowest for items in the middle. The parabolic recall rate is a combination of two effects: the primacy effect which refers to the enhanced recall for items presented at the beginning of the list, and the recency effect referring to the enhanced recall for items presented at the end of this list (Murdock, 1962). Previous research has shown that these separate effects are modulated by the stimulus modality. Recency effects, recalling items presented at the end, are stronger in auditory stimulus presentation compared to visual stimulus presentations (Crowder, 1986; Crowder & Morton, 1969; Engle & Mobley, 1976). Conversely, a stronger primacy effect was found when recalling information that was presented visually (Beaman, 2002). Crowder and Morton (1969) posit that the auditory recency effect is due to more elaborated processing by the aforementioned precategorical acoustic storage. The reasoning for the visual primacy effect is unclear, but it may be due to information being dual-coded, made of a subvocalized code and the visual code, leading to a more rich and durable representation of information (Paivio, 1990; Snodgrass et al., 1974). In the context of the base-rate neglect task, it may be the case that manipulating the order of acquisition for base rate information and vignette description could show an advantage for one
presentation modality over the other due to primacy/recency advantages exclusive to either modality.

**Present Research**

Previous research has shown that some types of working memory tasks invoke differences in performance between the auditory and visual modalities. However, it is unknown if presentation modality affects decision-making and reasoning tasks in the same manner. The current study posits that there are differences in performance on the base rate neglect task between auditory and visual presentation modalities. This study consists of two experiments. Experiment 1 investigates the performance of each modality on the base rate neglect task while manipulating the presentation order of information to see if a serial position-like effect exists in a decision-making and reasoning task. Experiment 2 tests the performance on the base rate neglect task while working memory is taxed through a secondary task.
Chapter 2:

Experiment 1: Base-“ic” Neglect Differences Between Modalities
Experiment 1

Modality differences exist in a variety of working memory tasks. However, there has not been a consideration of presentation modality affecting performance on decision-making and reasoning tasks, specifically the base rate neglect task. Using the base rate neglect task over other decision making and reasoning tasks allows for the manipulation of several variables making the experiment more sensitive to any differences between modalities.

In the present study, subjects performed a base rate neglect task presented in either the auditory or visual modality. The base rate neglect problems consisted of two parts, the base rate information (e.g., 997 engineers, 3 lawyers) and the vignette description of an individual. The order of the base rate and vignette description (base-rate information then vignette description; or vignette description then base-rate information), and base-rate problem congruency (congruent and incongruent) were counter-balanced across trials with an equal number for each order. All base-rate problems were adapted from Thompson et al. (2011). Following the presentation of the base-rate problem, subjects input their probability estimates using a sliding scale.

Hypotheses

Based on the results from previous base rate neglect literature in tandem with working memory, decision-making, and reasoning literature this study is designed to test the following hypotheses:

Hypothesis A

I expect that subjects in the visual modality group will have higher probability estimation accuracy than subjects in the auditory modality group. The reasoning for this hypothesis is based on the advantage of numerical recall that the visual group has over the auditory group, giving an advantage in recalling the base rate ratio when making probability judgements (Klingner et al.,
In addition, evidence suggests that working memory load is reduced in more difficult cognitive tasks for the visual modality compared to the auditory modality (Klingner et al., 2011). The more efficient use of working memory capacity will allow the visual group to have more working memory resources available and an increased capacity for conflict detection and the engagement of analytical thinking.

**Hypothesis B**

Previous research suggests a recency effect for auditory modality presentation, and a primacy effect for visual modality presentation (Conway & Christiansen, 2005). With this in mind, either one or both of these effects will be present in the base rate neglect task. I expect that on trials where the base rate information is presented first, the visual modality will have a higher accuracy in probability estimation. The dual-coded base rate information will be more durable and less prone to decay over the presentation period of the personality description. The recency effect would suggest that the auditory group will have higher accuracy than the visual group when the base rate information is presented after the personality description.

**Hypothesis C**

I expect to replicate results from previous research (Pennycook et al., 2015) showing that trials where the base rate information is presented before the vignette description will yield lower probability estimation accuracy than those trials where the base rate information is presented after the vignette description. This effect is due to the base rate information being more readily accessible as it is the most recent bit of stimuli prior to subjects’ probability estimation.

**Hypothesis D**

I expect to replicate results from previous research on congruency. That is, estimation accuracy for congruent problems will be higher than incongruent problems (Pennycook et al.,
The lack of conflict in reasoning in congruent problems, due to the analogous pieces of information, reinforces subject responses that match the base rate. In incongruent problems, due to the conflict in reasoning that occurs, a heuristic-based approach is used, leading to a reliance on the personality description when making judgements, and decreasing accuracy.

**Method**

**Subjects**

The study consisted of 148 subjects (51 Male, 94 Female, 3 Other, $M_{age} = 28.8$) collected from Prolific (www.prolific.co). All subjects were pre-screened on three parameters as all subjects had: English as their first language, normal or corrected to normal vision, and no hearing difficulties. Subjects were compensated £2.00 upon completion of the experiment.

**Apparatus**

The task was performed on each subject’s home computer using any internet browser of their choice except for Safari. The experiment was built using the python-based experiment software Psychopy (v2020.2.5). Using the built-in translating feature of Psychopy, it was converted into a JavaScript version of the experiment to run via internet browser. Experiment running, and data collection of the internet-browser based version of the experiment was hosted by Pavlovia (www.pavlovia.org).

**Stimuli**

In the auditory group, the base rate problems were recited by a male voice with a white fixation cross (0.1 x 0.1 normalized units$^1$) and a grey background. The presentation length for

---

$^1$ Normalized (‘norm’) units: the window scaling for x and y range from -1 to +1. The top edge of, and the right edge of, the window are +1. Whereas the bottom edge of, and left edge of, the window are -1
each auditory description was between 25-26 seconds. The presentation of base rate problems for the visual condition was presented in a rapid serial visual presentation format (RSVP). Each word was presented in white text (0.1 height norm units) in the centre of the screen over a grey background. The presentation duration for each word was the same as the auditory recitation except each word had a minimum presentation time of 350ms to ensure proper comprehension of the vignettes.

**Procedure**

Subjects were collected on Prolific (www.prolific.co). Testing was done in alternating groups of visual or auditory modality. The overall procedure and problems were the same for both groups (see Appendix A for the base rate descriptions). The main difference was the modality in which the problems were presented. Both subject groups were first presented with a screen with written instructions and examples of a typical trial for their respective task. The full task instructions given to subjects are in Appendix B. Subjects then performed two practice trials in their respective modality. One practice trial was Base Rate First/Congruent trial type and the other was Base Rate Last/Incongruent. The order of these two practice trials were randomized across subjects.

In both tasks, 16 Base Rate descriptions were presented in a random order. For the full list of the base rate problems and probability prompts refer to Appendix A. Each individual trial started with a central fixation cross presented for 1000ms, followed by the base rate problem. Each base rate problem contained a description of a sample population of 1000 with two separate groups (e.g., 995 lawyers and 5 engineers), and a stereotypical description of an individual randomly selected from the 1000 people. All sample populations used extreme base rate ratios with three different ratios (995:5, 996:4, and 997:3). Base rate problems were adapted from
Thompson et al. (2011) to ensure base rate scenarios were between 49 to 52 words in length. In Congruent trials the description of the individual matches stereotypical features of the higher population group. In Incongruent trials the description of the individual matches stereotypical features of the smaller population group. In addition, for half of the base rate problems the base rate information (e.g., In a sample of 1000 people there is 995 lawyers and 5 engineers) was presented before the vignette, and the other half was presented after the vignette. Of the 16 base rate problems, each subject would be tested with 4 Base Rate First/Congruent, 4 Base Rate First/Incongruent, 4 Base Rate Last/Congruent, and 4 Base Rate Last/Incongruent trial types.

After the presentation of the problem, subjects were required to estimate the probability that the person described in the problem belonged to one of the two groups. The response screen consisted of the written text asking “What is the probability that [person] is a [group]” over a sliding scale from 0 to 100 with tick markers and numbering in intervals of 10. Subjects were equally divided into two groups to counterbalance which specific population group would be probed of the two for each trial. Subjects reported their probability estimate by marking a point on the scale using the computer mouse. They completed their response by pressing the space-bar on the computer keyboard to submit their estimate. Response times for their probability estimations were also recorded, starting from the onset of the prompt for probability estimation until subjects press the “space-bar” to submit their answer. Upon pressing the space-bar the experiment advanced to the next trial.

The general experiment paradigm for the visual task is shown in Figure 1 and for the auditory task is shown in Figure 2. The only difference from the visual modality group being the presentation of the base rate descriptions, which are presented auditory with a male voice with a
central fixation cross in the centre of the screen. Each auditory presentation of the base rate problems is 25s in length.
Figure 1. General experimental paradigm for the visual group in Experiment 1. Trials begin with a fixation-cross presented in the centre of the screen for 1000ms. After which the base rate problem is presented in an RSVP format, where each word is presented individually in sequence (figure diagram is an example of the display). After the presentation of the base rate problem, subjects are asked to estimate the probability that the person described is part of the group prompted using the sliding scale. After subjects input their probability estimation, the next trial begins.
Figure 2. General experimental paradigm for the auditory group in Experiment 1. Trials begin with a fixation-cross presented in the centre of the screen. After 1000ms, an auditory recording of a male voice is played, presenting the base rate problem. After the presentation of the base rate problem, subjects are asked to estimate the probability that the person described is part of the group prompted. After subjects input their probability estimation, the next trial begins.
After subjects completed the experiment, subjects were asked to complete a short two question post-experiment survey asking subjects to assess the difficulty of the task. The questions asked were “How difficult did you feel the task was overall?” and “How difficult was remembering the descriptions when making judgements?”, in that order. Each question was presented over a sliding Likert scale from 1 to 5 (1 being “very easy” and 5 being “very difficult”) where subjects would submit their answers using the sliding scale.

Results

The mean response time for probability judgements was 8.76 seconds across both modalities. Seven subjects with mean response times for probability estimates greater than 2.5SD were removed from analyses. This leaves a total of 141 subjects of the original 148. A 2 (Congruency[Congruent, Incongruent]) x 2 (Base Rate Position[First, Last]) x 2 (Modality[Audio, Visual]) repeated measures ANOVA with modality being a between group variable. This mixed model ANOVA was used to analyze the dependent variables of base rate neglect accuracy and response time. All post-hoc analyses were done using Tukey post-hoc analyses, unless otherwise stated. There were 71 subjects in the Auditory group, and 70 subjects in the Visual group.

Base Rate Neglect Accuracy

Probability estimations were standardized so that a higher number (closer to 100) represents a closer accuracy to the correct probability estimation. This was done for trials where the correct estimated probability would be close to 0 (the smaller population group). In these trials, subjects’ probability estimations were subtracted from 100. In all other trials, subject responses were left untouched.
**Congruency.** The main effect of congruency was significant, $F_{(1, 139)} = 194.547, p < 0.001, \eta_p^2 = 0.583$. Post-hoc analyses, $p < 0.001$, showed that the mean probability estimation accuracy for congruent trials ($M = 81.5, SD = 13.7$) was more accurate than incongruent trials ($M = 53, SD = 27.7$).

**Base Rate Position.** The main effect of base rate position was significant, $F_{(1, 139)} = 31.714, p < 0.001, \eta_p^2 = 0.186$. Post-hoc analyses were significant, $p < 0.001$, indicating that trials where the base rate information was presented first ($M = 63.7, SD = 19.3$) were lower in base rate probability accuracy than on trials where the base rate information was presented last ($M = 70.9, SD = 19.9$).

**Modality.** There was no main effect of modality, $F_{(1, 139)} = 1.17, p = 0.28, \eta_p^2 = 0.008$, indicating probability estimations were not significantly different between the visual modality ($M = 68.9, SD = 18.3$) and the auditory modality ($M = 65.6, SD = 17.9$).

**Interaction effects.** There was a significant interaction effect between Congruency and Base Rate Position, $F_{(1, 139)} = 18.582, p < 0.001, \eta_p^2 = 0.118$. Post-hoc analyses, $p < 0.001$, show that in incongruent trials the presentation of base rate information first ($M = 47.1, SD = 29.9$) yielded lower accuracy compared to being presented last ($M = 58.9, SD = 29.8$) but not in congruent trials ($M_{diff} = -2.58, t(139) = -1.84, p = 0.257$). Refer to Figure 3 for visualization.

There were no interaction effects found for Modality x Congruency ($F_{(1, 139)} = 2.199, p = 0.140, \eta_p^2 = 0.016$) or Modality x Base Rate Position ($F_{(1, 139)} = 0.225, p = 0.636, \eta_p^2 = 0.002$). In addition, there was no significant interaction for Modality x Base Rate Position x Congruency ($F_{(1, 139)} = 0.434, p = 0.511, \eta_p^2 = 0.003$).
Figure 3. Mean probability estimates for Experiment 1 as a function of base rate ratio position x base rate problem congruency. Error bars represent standard error.
Reaction Times

Reaction times were measured from the onset of the prompt asking for subject’s probability judgements to when they pressed the “space bar” to submit their answer and continue to the next trial.

There was no main effect for reaction time on Congruency ($F_{(1, 139)} = 0.136, p = 0.712, \eta_p^2 = 0.001$), Base Rate Position ($F_{(1, 139)} = 0.688, p = 0.408, \eta_p^2 = 0.005$), and Modality ($F_{(1, 139)} = 3.53e^{-4}, p = 0.985, \eta_p^2 = 0.000$). In addition, post-hoc analyses showed no significant interactions for Congruency x Base Rate Position ($F_{(1, 139)} = 1.685, p = 0.196, \eta_p^2 = 0.012$), Base Rate Position x Modality ($F_{(1, 139)} = 0.015, p = 0.902, \eta_p^2 = 0.000$), Congruency x Modality ($F_{(1, 139)} = 3.675, p = 0.057, \eta_p^2 = 0.026$). Finally, there was no Congruency x Base Rate Position x Modality interaction, $F_{(1, 139)} = 0.636, p = 0.427, \eta_p^2 = 0.005$.

Post Experiment Questionnaire

The responses for the post-experiment questionnaire were on a 5-point Likert scale with “1” labelled as “very easy” and “5” as “very difficult”. Analyses were done using an independent t-test with Modality as the between group comparison. Two subjects did not provide post-questionnaire feedback, leaving 137 subject responses.

**Difficulty.** The responses to the question “How difficult did you find the task overall?” were not found to be significantly different between modalities, $t(137) = -0.6953, p = 0.488, d = -0.118$.

**Remembering.** The responses to the question “How difficult was it to remember descriptions when making judgements?” were not found to be significantly different between modalities, $t(137) = 0.0713, p = 0.943, d = 0.012$. 


Discussion

While some of the results from Experiment 1 were as expected, the main comparison between modality performance did not yield a significant difference between modality groups. The probability estimates when manipulating congruency and base rate position are consistent with previous literature (Thompson et al., 2011). Congruent problems resulted in higher base rate probability estimation accuracy compared to that of incongruent problems. This is due to a lack of conflict in reasoning in congruent problems, whereas in incongruent problems the information conflict between the base rate information and the personality description leads subjects to rely more on the personality description in their probability estimations. The base rate position effect was also significant, showing that when the base rate information is presented last, subjects have higher probability accuracy. This is due to the base rate information being more salient, as it is the last piece of information presented prior to subjects’ probability estimations.

When looking at the difference in performance between the auditory and visual modality, there were no significant differences in performance between them, failing to support hypothesis A. However, the results supported hypotheses C and D, giving confidence that subjects performed the task correctly. Before making any conclusions based on these results, other explanations need to be addressed. Primarily, it may be that the task itself is not sensitive enough to detect differences in performance, or that the task itself does not tax working memory enough for modality differences to occur.
Chapter 3:

Experiment 2: Testing Modalities’ Loaded Performance
Experiment 2

The major findings from Experiment 1 yielded no significant differences in base rate probability estimates between the auditory and visual tasks contrary to our main hypothesis of interest. To further investigate potential differences in modality performance on the base rate neglect task, a second experiment is proposed. In Experiment 1, it may have been that the base rate neglect task itself was not sensitive enough to detect any differences in modality performance. To increase the sensitivity of the base rate neglect task, it may be the case that additional working memory load is needed to see performance differences between modalities. Should one modality be more impacted by higher working memory load, that modality would be more susceptible to base rate neglect.

To explore this possibility Experiment 2 tested subjects in the same base rate neglect task in either one of the modality conditions, as in Experiment 1, but this time a secondary concurrent task was introduced to manipulate subjects’ memory load. This memory load task consisted of a different simple letter array of either three or seven randomly determined letters presented at the beginning of each trial. Subjects were instructed to memorize the array and hold it in memory during the presentation of the base rate neglect description. After subjects made their base rate estimate they would be presented with either a new letter or one of the letters from the array presented at the beginning of the trial. Subjects made a two-alternative forced choice to report whether that test letter was in that array.

The rationale for using a letter array for the memory load task was to engage as much as possible the same processes utilized in the auditory and visual processing of the base rate problems. Using a numerical or spatial task may give too much of an advantage to the visual modality group. Similarly, the task cannot be so overly difficult that it would introduce floor
effects of performance for both modalities across both tasks. Memory load tasks are also an effective method to induce working memory demands during the base rate task. Other literature has shown that as memory load increases, so too does pupil dilation indicating higher levels of cognitive resource strain (Granholm et al., 1996; Peavler, 1974; Kahneman & Beatty, 1966; for a review see van der Wel & Steenbergen, 2018).

The decision to use the sizes of three and seven letter arrays was done to create an adequate comparison of low (three letters) and high (seven letters) memory loads. Similar studies using arrays have indicated that low load arrays are typically four or fewer items, moderate load arrays are between five and nine items, and any arrays over nine items overload working memory (Peavler 1974; Cabestrero et al., 2009). However, when using nine or more items as the “high” memory load level, research shows that subjects’ pupil dilation decreases, indicating working memory strain decreases due to an overload of working memory (Peavler, 1974).

The base rate problems used were the same as that in Experiment 1 except here base rate first ordering was used for each trial. This was to make it more comparable to tasks in previous base rate neglect literature, as Experiment 1 showed the base rate presentation order made no difference in either modality’s performance.

The presentation order of the two tasks (secondary task load, the base rate neglect task presentation, base rate probability prompt, secondary task prompt) was done to ensure that the secondary task load would be kept in working memory during both the presentation of the base rate description and during subjects’ probability estimations.

Hypotheses

Given the results from Experiment 1 and the addition of a secondary working memory task the focus of Experiment 2 is to test three hypotheses:
Hypothesis A

Any differences in working memory will be shown in performance on the base rate neglect task due to the addition of cognitive load by the secondary task. To ensure subjects dedicate working memory resources primarily to the secondary task, an added monetary incentive based on secondary task performance is given. This will allow for any advantages in working memory processing or capacity to be apparent between the two modalities. Specifically, due to the same reasons outlined in Experiment 1, Hypothesis A, the visual modality will have greater base-rate neglect accuracy.

Hypothesis B

In the secondary task, I expect the performance between modalities to be similar. This is due to the added incentive for performance for the secondary memory load task, as subjects in either modality will be equally motivated to perform well.

Hypothesis C

As per the hypotheses for Experiment 1, I expect to see similar findings for congruency, with higher base rate accuracy for congruent problems than incongruent ones.

Method

Subjects

This experiment consisted of 148 subjects (69 male, 76 female, 3 other, \( M_{age} = 33.65 \)) collected from Prolific (www.prolific.co). However, six subjects did not provide an age in part of their demographics and were removed from future analyses. Subjects were pre-screened on the same parameters as Experiment 1. Subjects were compensated £2.13 upon completion of the experiment. In addition, an extra incentive of £1.00 was given to those whose accuracy on the memory load task was higher than 81.25% (13 correct probes of the possible 16).
**Apparatus**

This was the same as outlined in Experiment 1. The task was performed on each subject’s home computer using any internet browser of their choice except for Safari. The experiment was built using the python-based experiment software Psychopy (v2020.2.5). Using the built-in translating feature of Psychopy, it was converted into a JavaScript version of the experiment to run via internet browser. Experiment running and data collection the internet-browser based version of the experiment was hosted by Pavlovia (www.pavlovia.org).

**Stimuli**

The stimuli in the base rate task were the same as in Experiment 1 except in all base rate descriptions the base rate information was presented before the vignette description. In the auditory group, the base rate problems were recited by a male voice with a white fixation cross (0.1 x 0.1 norm units) and a grey background. The presentation length for each auditory description was between 25-26 seconds. The presentation of the base rate problems for the visual condition was presented in an RSVP format. Each word was presented in white text (0.1 height norm units) in the centre of the screen over a grey background. The presentation duration for each word was the same as the auditory recitation except each word had a minimum presentation time of 350ms to ensure proper comprehension of the vignettes.

The memory load task consisted of a white fixation cross (0.1 x 0.1 norm units) presented on a grey background for 1500ms followed by the presentation of an array of white capitalized letters for 3500ms followed by another fixation cross for 1000ms preceding the base rate descriptions in their respective modality presentation. The letter array consisted of either three or seven letters. Letters were randomly determined for each trial. No letters were repeated within the same array. Each individual base rate problem had its own memory load and prompt which
was the same between subjects and across modalities. All base rate problems were adapted from Thompson et al. (2011); see Appendix C for the full list of base rate problems, and their respective memory loads and prompts.

**Procedure**

The general procedure for Experiment 2 was similar to that of Experiment 1. With the exception of the addition of the memory load task. Subjects were first presented with a screen with written instructions about the experiment. The instructions were similar to those in Experiment 1, except subjects were also given instructions on how to perform the memory load task and told they earn an extra £1.00 if they obtained 81.25% accuracy in the memory load task. The reason for the incentive to perform well on the memory load task was to increase our chances of the subjects’ compliance in performing the load task. After the instruction screens, subjects completed two practice trials. The following practice problems were the same as that in Experiment 1 except for the addition of the memory load task. Full instruction details can be found in Appendix D.

Trials were identical as in Experiment 1 with only two exceptions: base rate problems were only in the base rate first format and the memory load task. In a typical trial, subjects were presented with the fixation cross (1500ms) followed by the letter array (3500ms). Following the letter array, subjects were presented with the base rate description in their respective modality. Subjects are prompted about the base rate probability in the same manner as in Experiment 1. After subjects submitted their probability estimate, they were then presented with a new screen consisting of a letter that was either one of the letters used in the letter array at the beginning of the trial (old letter) or a letter not in that letter array (new letter). The screen also consisted of text with the question: “Did it contain a ___?” with the blank being either the old or new letter.
Subjects reported whether they had seen the letter via a two-alternative forced choice response by pressing the keyboard button “1” for yes or “2” for no. Upon responding the experiment moved on to the next trial. See Figure 4 for the visual format of the general experiment paradigm for Experiment 2, and Figure 5 for the auditory format.
Figure 4. The general experimental paradigm for the visual presentation of Experiment 2. Trials begin with a fixation-cross presented in the centre of the screen for 1500ms. After which the letter array corresponding to the specific base rate problem is presented (either 3 or 7 items) for 3500ms. A fixation cross is then presented for 1000ms prior to the base rate problem. The base rate problem is presented in an RSVP format, where each word is presented individually in sequence (the figure diagram is an example of the display). After the presentation of the base rate problem, subjects are asked to estimate the probability that the person described is part of the group prompted using the sliding scale. Following subjects’ estimation, they are asked if the letter array at the beginning of the trial contained the prompted letter. After subjects answer the two-alternative forced-choice input, the next trial begins.
Figure 5. The general experimental paradigm for the auditory presentation of Experiment 2. Trials begin with a fixation-cross presented in the centre of the screen for 15000ms. After which the letter array corresponding to the specific base rate problem is presented (either 3 or 7 items) for 3500ms. A fixation cross is then presented for 1000ms prior to the base rate problem. Then, an auditory recording of a male voice is played, presenting the base rate problem. After the presentation of the base rate problem, subjects are asked to estimate the probability that the person described is part of the group prompted using the sliding scale. Following subjects’ estimation, they are asked if the letter array at the beginning of the trial contained the prompted letter. After subjects answer the two-alternative forced-choice input, the next trial begins.
After 16 trials, subjects are given the same post-experiment questionnaire from Experiment 1 with the addition of two new questions. Subjects were asked “How difficult was it to remember 3 letter strings?” and “How difficult was it to remember 7 letter strings?”.

Results

The mean response time for base rate probability judgement was 9.16 seconds across both modalities. Three subjects with a mean response time greater than 2.5SD of the mean were removed, leaving 139 subjects for analyses. A 2 (Congruency[Congruent, Incongruent]) x 2 (Array Size[3, 7]) x 2 (Modality[Audio, Visual]) repeated measures ANOVA with modality being a between group variable. This mixed model ANOVA was used to analyze the dependent variables of base rate neglect accuracy, memory load accuracy, and response time for each. All post-hoc analyses were done using Tukey post-hoc analyses, unless otherwise stated. There were 71 subjects in the Visual group, and 68 subjects in the Auditory group.

**Base Rate Neglect Accuracy**

**Congruency.** There was a significant main effect of congruency, $F(1, 137) = 180.418, p < 0.001, \eta^2_p = 0.568$. Post-hoc analyses, $p < 0.001$, showed that congruent problems ($M = 83.3, SD = 13.529$) had a higher accuracy than incongruent problems ($M = 50.4, SD = 28.307$).

**Array Size.** There was a significant main effect of array size on base rate accuracy, $F(1, 137) = 8.734, p = 0.004, \eta^2_p = 0.060$. Post-hoc analyses, $p = 0.004$, revealed that subjects were less accurate in their base rate estimations for the 3 array size ($M = 65.4, SD = 17.921$) than the 7 array size ($M = 68.3, SD = 17.771$).

**Modality.** There was no significant main effect of modality, $F(1, 137) = 1.12, p = 0.293, \eta^2_p = 0.008$, indicating no difference in performance between the visual modality ($M = 68.4, SD = 18.364$) and the auditory modality ($M = 65.3, SD = 15.096$).
**Interaction Effects.** There was an interaction effect for Array Size x Congruency, $F_{(1, 137)} = 28.922, p < 0.001, \eta_p^2 = 0.174$. Post-hoc analyses showed, $p < 0.001$, in incongruent trials, subjects’ base rate accuracy was lower with 3 array size ($M = 45.9, SD = 30.6$) than 7 array size ($M = 54.9, SD = 29$). However, no differences between array sizes were found in congruent trials, ($M_{diff} = 3.03, t(137) = 2.17, p = 0.137$). Refer to Figure 6 for visualization. There was no significant interaction effect for Modality x Array Size ($F_{(1, 137)} = 1.166, p = 0.282, \eta_p^2 = 0.008$) or Modality x Congruency ($F_{(1, 137)} = 0.066, p = 0.798, \eta_p^2 = 0.000$). There also was no Array Size x Congruency x Modality interaction effect, $F_{(1, 137)} = 2.01, p = 0.159, \eta_p^2 = 0.014$. 
Figure 6. Mean probability estimates for Experiment 2 as a function of memory load size x base rate problem congruency. Error bars represent standard error.
**Reaction Time**

**Congruency.** There was a main effect of Congruency for reaction time, $F_{(1, 137)} = 5.965$, $p = 0.016$, $\eta^2_p = 0.042$. Post-hoc analyses were significant, $p = 0.016$, indicating that the subjects were faster to respond to the base rate problems in congruent trials ($M = 8.54s$, $SD = 3.669$) than incongruent trials ($M = 9.1s$, $SD = 4.025$).

There were no main effects of Array Size, $F_{(1, 137)} = 0.646$, $p = 0.423$, $\eta^2_p = 0.005$, or Modality, $F_{(1, 137)} = 0.039$, $p = 0.844$, $\eta^2_p = 0.000$. In addition, there was no interaction effects for Array Size x Congruency ($F_{(1, 137)} = 0.707$, $p = 0.402$, $\eta^2_p = 0.005$), Array Size x Modality ($F_{(1, 137)} = 2.961$, $p = 0.088$, $\eta^2_p = 0.021$), Congruency x Modality ($F_{(1, 137)} = 0.852$, $p = 0.358$, $\eta^2_p = 0.006$), or Array Size x Congruency x Modality ($F_{(1, 137)} = 0.135$, $p = 0.714$, $\eta^2_p = 0.001$).

**Memory Load Accuracy**

Memory load accuracy is represented as a percentage of correct memory load probes.

**Array Size.** There was a main effect of array size for memory load accuracy, $F_{(1, 137)} = 144.918$, $p < 0.001$, $\eta^2_p = 0.514$. Post-hoc analyses revealed, $p < 0.001$, that subjects were more accurate when array sizes were 3 letters long ($M = 87.5$, $SD = 11.5$) than 7 letters long ($M = 72.6$, $SD = 19.6$).

**Congruency.** There was a main effect of congruency for memory load accuracy, $F_{(1, 137)} = 107.4$, $p < 0.001$, $\eta^2_p = 0.439$. Post-hoc analyses revealed, $p < 0.001$, that subjects were less accurate in the congruent condition ($M = 74.0$, $SD = 19.6$) than the incongruent condition ($M = 86.1$, $SD = 11.8$).

**Modality.** There was a main effect of modality on memory load accuracy, $F_{(1, 137)} = 214$, $p < 0.001$, $\eta^2_p = 0.609$. Post-hoc analyses revealed, $p < 0.001$, that subjects in the visual
condition ($M=90.5$, $SD = 9.26$) were more accurate than the auditory group ($M = 69.6$, $SD = 7.47$).

**Interaction effects.** There were several interaction effects present in the post-hoc analyses. There was a significant interaction effect for Array Size x Modality, $F_{(1, 137)} = 64.773$, $p < 0.001$, $\eta^2_p = 0.321$. Post-hoc analyses show ($M_{\text{diff}} = 24.82$, $t(137) = 14.05$, $p < 0.001$) that there was a greater decrease in accuracy for the auditory group when comparing an array size of 3 ($M = 82$, $SD = 9.51$) to an array size of 7 ($M = 57.2$, $SD = 11.5$) than the decrease in the visual group ($M_{\text{diff}} = 4.93$, $t(137) = 2.85$, $p = 0.026$) between array size of 3 ($M = 93$, $SD = 10.7$) and 7 ($M = 88$, $SD = 12.6$). Refer to Figure 7 for visualization.

In addition, a significant interaction effect of Congruency x Modality, $F_{(1, 137)} = 101.206$, $p < 0.001$, $\eta^2_p = 0.425$, was found. Post-hoc analyses, $p < 0.001$ showed the auditory group had a decrease in accuracy in the congruent condition ($M = 57.7$, $SD = 10.1$) compared to the incongruent condition ($M = 81.4$, $SD = 10.9$). Whereas the visual group showed no difference in performance between congruent and incongruent conditions ($M_{\text{diff}} = -0.352$, $t(137) = -0.217$, $p = 0.996$). Refer to Figure 8 for visualization.

There were no significant interaction effects found for Array Size x Congruency ($F_{(1, 137)} = 0.698$, $p = 0.405$, $\eta^2_p = 0.005$), or Array Size x Congruency x Modality ($F_{(1, 137)} = 0.294$, $p = 0.589$, $\eta^2_p = 0.002$).
Figure 7. Mean memory load accuracy for Experiment 2 as a function of Array Size x Modality. Error bars represent standard error.
Figure 8. Mean memory load accuracy for Experiment 2 as a function of congruency x modality. Error bars represent standard error.
Cowan’s K

To further examine the effects on memory load performance the subjects’ responses to the memory load probe were transformed into Cowan’s K (Cowan, 2000). Cowan’s K is a value denoting the storage capacity of working memory as the number of items that can be stored (Cowan, 2000). Cowan’s K is calculated using the formulae \( K = N \times (H-FA) \), where \( N \) represents the array size (3 or 7), \( H \) is the hit rate or number of trials correct when the probed letter was a part of the original letter array, and \( FA \) is the false alarm rate or number of incorrect trials when the probed letter was not part of the initial letter array.

**Array Size.** There was a main effect of array size for Cowan’s K, \( F(1, 137) = 39.9, p < 0.001, \eta^2_p = 0.226 \). Post-hoc analyses revealed, \( p < 0.001 \), that the array size of 3 storage capacity (\( M = 4.50, SD = 1.216 \)) was less than the array size of 7 (\( M = 6.33, SD = 3.376 \)).

**Congruency.** There was a main effect of congruency for Cowan’s K, \( F(1, 137) = 91.9, p < 0.001, \eta^2_p = 0.401 \). Post-hoc analyses revealed, \( p < 0.001 \), that Congruent storage capacity (\( M = 4.17, SD = 2.328 \)) was less than Incongruent (\( M = 6.66, SD = 2.514 \)).

**Modality.** There was a main effect of modality for Cowan’s K, \( F(1, 137) = 244, p < 0.001, \eta^2_p = 0.641 \). Post-hoc analyses revealed, \( p < 0.001 \), that visual modality subjects’ storage capacity (\( M = 7.90, SD = 2.618 \)) was higher than the subjects in the auditory condition (\( M = 2.92, SD = 2.688 \)).

**Interaction Effects.** There were several interaction effects present in the post-hoc analyses. There was a significant interaction effect for Array Size x Congruency, \( F(1, 137) = 19.2, p < 0.001, \eta^2_p = 0.123 \). Post-hoc analyses, \( p < 0.001 \), showed the Incongruent condition had decreased storage capacity from the 7 array size (\( M = 8.15, SD = 5.32 \)) to the 3 array size (\( M = 5.16, SD = 1.534 \)). Whereas the Congruent condition showed no difference in storage capacity.
(\(M_{\text{diff}} = -0.668, t(137) = -1.76, p = 0.297\)) between array sizes of 3 (\(M = 3.86, SD = 2.20\)) and 7 (\(M = 4.64, SD = 7.38\)).

There was a significant interaction effect for Array Size x Modality, \(F_{(1, 137)} = 159.7, p < 0.001, \eta^2_p = 0.538\). Post-hoc analyses, \(p < 0.001\), showed that for the visual condition, there was a decrease in storage capacity from the 7 array size (\(M = 10.65, SD = 3.52\)) to the 3 array size (\(M = 5.16, SD = 1.286\)). Post-hoc analyses showed, \(p < 0.001\), that in the auditory group there was an increase in storage capacity from the 7 array size (\(M = 2.01, SD = 3.364\)) to 3 array size (\(M = 3.84, SD = 1.22\)).

There was a significant interaction effect for Congruency x Modality, \(F_{(1, 137)} = 84.7, p < 0.001, \eta^2_p = 0.382\). Post-hoc analyses showed that for the auditory group, \(p < 0.001\), Congruent conditions had a lower storage capacity (\(M = 0.485, SD = 2.33\)) than Incongruent conditions (\(M = 5.36, SD = 2.54\)). In the visual group, there was no difference in storage capacity between Congruent or Incongruent trials (\(M_{\text{diff}} = -0.098, t(137) = -0.272, p = 0.993\)).

There was a significant interaction effect for Array Size x Congruency x Modality, \(F_{(1, 137)} = 16.1, p < 0.001, \eta^2_p = 0.105\). Comparing the visual modality on both levels of array size between the congruent and incongruent problems, there was no difference in storage capacity for subjects in the visual modality. Subjects in the auditory modality showed a difference in storage capacity on array size depending on congruency. In the Incongruent conditions, there was no storage capacity difference between array sizes of 3 (\(M = 5.16, SD = 1.54\)) and 7 (\(M = 5.56, SD = 4.66\)). In the congruent condition, there was a significant difference, \(p < 0.001\), where the storage capacity for 7 item arrays (\(M = -1.54, SD = 4.205\)) was lower than 3 item arrays (\(M = 2.51, SD = 1.764\)). Refer to Figure 9 for visualization.
Figure 9. Cowan’s K values for Experiment 2 as a function of congruency x modality x array size. Error bars represent standard error.
Memory Load Reaction Time

Reaction time was measured from the onset of the memory load prompt screen until subjects answered.

Array Size. There was a main effect for array size on reaction time for answering the memory load prompt, $F_{(1, 137)} = 12.406, p < 0.001, \eta^2_p = 0.083$. Post-hoc analyses revealed that subjects took less time to respond when the array size was 3 ($M = 2.88s, SD = 1.936$) than when the array size was 7 ($M = 3.58s, SD = 2.719$).

Congruency. There was no main effect of congruency on memory load prompt reaction time, $F_{(1, 137)} = 0.09, p = 0.765, \eta^2_p = 0.001$.

Modality. There was no main effect of modality on memory load prompt reaction time, $F_{(1, 137)} = 0.543, p < 0.462, \eta^2_p = 0.004$.

Interaction Effects. There were no significant interaction effects for Array Size x Modality ($F_{(1, 137)} = 9.68e^5, p = 0.992, \eta^2_p = 0.000$), Congruency x Modality ($F_{(1, 137)} = 0.16, p = 0.689, \eta^2_p = 0.001$), or Array Size X Congruency ($F_{(1, 137)} = 2.716, p = 0.102, \eta^2_p = 0.019$).

However, there was a significant three-way interaction effect for Array Size x Congruency x Modality, $F_{(1, 137)} = 5.551, p = 0.02, \eta^2_p = 0.039$. The reaction times of the visual group on the item array sizes of 3 and 7 for congruent problems are the same as their respective array size in the incongruent condition. However, the auditory modality has similar reaction times in congruent problems between array sizes of 3 ($M = 3.41s, SD = 2.48$) and 7 ($M = 3.33s, SD = 2.33$). For incongruent problems, subjects were faster in responding for 3 item arrays ($M = 2.62s, SD = 1.84$) than 7 item arrays ($M = 4.08s, SD = 4.11$). Refer to Figure X for visualization.
Figure 10. Reaction time for responding to the memory load task (in seconds) for Experiment 2 as a function of congruency x modality x array size. Error bars represent standard error.
Post Experiment Questionnaire

The post-experiment experiment was the same as Experiment 1, with two additional questions. Analyses were done using independent t-tests with modality as the between group measure.

Difficulty. There was no significant difference between modalities in responses to the question “How difficult did you find the task overall?”, $t(137) = -1.138$, $p = 0.257$, $d = -0.193$.

Remembering. There was a significant difference between modalities in response to the question “How difficult was it to remember descriptions when making judgements?”, $t(137) = -2.026$, $p = 0.045$, $d = -0.344$. A comparison of means shows that the Auditory group experienced more difficulty remembering descriptions ($M = 2.90$, $SD = 1.019$) than the Visual group ($M = 2.54$, $SD = 1.06$).

Recognizing 3 Length Arrays. There was no significant difference in response between modalities when asked “How difficult was it to remember the 3 letter arrays?”, $t(137) = 0.469$, $p = 0.640$, $d = 0.08$.

Recognizing 7 Length Arrays. There was no significant difference in response between modalities when asked “How difficult was it to remember the 7 letter arrays?”, $t(137) = 0.033$, $p = 0.974$, $d = 0.006$.

Discussion

The results of Experiment 2 introduced some interesting findings. The effect of congruency was similar to that of Experiment 1, where congruent problems had a higher base rate accuracy than incongruent problems. However, there was no difference between modalities in base rate accuracy. A comparison between the results of Experiment 2 and base rate first
problems from Experiment 1 showed there was no change in base rate accuracy when adding the memory task.

The results of the memory load task showed that subjects had higher recognition rates for three item arrays than seven item arrays. In addition, there was an unexpected difference in performance between the modalities, where the visual modality subjects were more accurate in their recognition rates than the auditory modality subjects. Converting the results on the memory load task to Cowan’s K, also showed a distinct difference in memory storage between modalities, with subjects in the auditory modality exhibiting a reduction in memory storage compared to subjects in the visual modality.

While the outcomes of Experiment 2 were unexpected, it does provide an amount of insight into the possible working memory differences between the auditory and visual modality.

**Comparison Between Experiment 1 and Experiment 2**

To determine if the memory task impacted base rate neglect task performance in Experiment 2, a comparison between Experiment 1 and Experiment 2 was conducted using a 2 (Congruency [Congruent, Incongruent]) x 2 (Experiment [1, 2]) x 2 (Modality [Audio, Visual]) repeated measures ANOVA on base rate neglect accuracy and reaction time. To properly compare accuracy between the two experiments, the data used from Experiment 1 were only from trials where the base rate information was presented first, the same as the base rate presentation ordering in Experiment 2. All post-hoc analyses were done using Tukey post-hoc analyses, unless otherwise stated.

**Base Rate Neglect Accuracy**

**Experiment.** There was no significant difference in base rate neglect accuracy, $F_{(1, 276)} = 2.122, p = 0.146, \eta_p^2 = 0.008$, between Experiment 1 and Experiment 2.
**Reaction Time**

**Experiment.** There was a significant main effect of Experiment on reaction time, $F_{(1, 276)} = 6.404$, $p = 0.012$, $\eta^2_p = 0.023$. Post-hoc analyses show the subjects took longer to respond in Experiment 2 ($M = 8.82s$, $SD = 3.45$) than Experiment 1 ($M = 7.78$, $SD = 3.44$).

**Post-Experiment Questionnaire**

A comparison of responses to the post-experiment survey between Experiment 1 and Experiment 2 was done to determine if the addition of a memory task impacted the subjects’ experiences of the base rate neglect task. This was done using a 2 (Modality [Audio, Visual]) x 2 (Experiment [1, 2]) ANOVA on the post-experiment survey responses of the questions regarding task difficulty and difficulty recalling base rate descriptions.

**Difficulty.** There was a significant difference in subject responses between modalities when asked “How difficult did you find the task overall?”, $F_{(1, 274)} = 22.867$, $p = 0.001$, $\eta^2_p = 0.077$. Subjects in the auditory modality reported higher perceived task difficulty in Experiment 2 ($M = 3.12$, $SD = 1.45$) than Experiment 1 ($M = 2.32$, $SD = 1.44$). Subjects in the visual modality did not differ in their ratings of task difficulty between Experiment 1 and 2 ($M_{diff} = -0.380$, $t(274) = -2.21$, $p = 0.124$).

**Remembering.** There was no significant difference in perceived difficulty when asked “How difficult was it to remember descriptions when making judgements?” between experiments ($F_{(1, 274)} = 4.71e^{-5}$, $p = 0.995$, $\eta^2_p = 0.000$). However, there is a main effect of modality ($F_{(1, 274)} = 4.423$, $p = 0.042$, $\eta^2_p = 0.015$), where the subjects in the auditory group reported more difficulty in remembering descriptions when making judgements ($M = 2.84$, $SD = 1.02$) than the visual group ($M = 2.59$, $SD = 1.02$).
Chapter 4:

Discussion
General Discussion

The present study was aimed at determining whether performance in decision-making and reasoning processes depends on the modality of presentation in the base rate neglect task. Previous literature has shown performance differences between stimulus modalities in a variety of other cognition tasks. To my knowledge, no study has investigated whether there are potential performance differences dependent on stimulus modality in a decision-making and reasoning task. The present study sought to find evidence of such differences using two versions of the base rate neglect task: an auditory version where the problems were presented orally via a pre-recorded audio presentation and a visual version where the problems were presented as text on a computer screen. In Experiment 1, the overall task performance in both groups was similar to previous base rate neglect studies, and the main results show no statistical difference in performance between the two modality groups. That is, accuracy in the visual and auditory groups was statistically the same. In light of these results, I explored the possibility that this lack of significant difference between modality groups might have been due to a lack of sensitivity in detecting differences between modalities in the experiment. Experiment 2 repeated Experiment 1 but this time a memory load task was added, thereby increasing sensitivity to working memory differences that exist between modalities. The reasoning for adding the memory load task is as follows: if either modality is affected by the additional working memory load induced by the memory load task, then any differences in cognitive processing would be apparent in the base rate neglect task. The results of Experiment 2 replicated the base rate neglect results of Experiment 1, including the lack of a significant main effect of modality group. However, and unexpectedly, the memory load task showed a distinct performance difference between the auditory and visual modality, where the visual modality was more accurate than the auditory
modality. This was an unexpected result, as the performance on the memory load task was expected to be statistically the same between modality groups due to the incentivization for performance, and indeed the memory load task was intended to reduce the processing capacity for the main base rate neglect task and reveal a potential difference in base rate neglect performance between modalities. The results of Experiment 2 found the opposite effect of differences between modality groups in the memory load task and not the base rate neglect task. In the following section, I will elaborate on the findings from each experiment, compare the two experiments, and discuss the limitations and potential future directions of further research.

**Experiment 1 Findings**

The primary objective of Experiment 1 was to elucidate potential performance differences due to stimulus presentation between the auditory and visual modality in a base rate neglect task. The results showed no performance differences between modalities on either base rate accuracy or reaction time in responding. However, the results were very consistent with typical results of previous base rate neglect studies in both base rate position and congruency (Thompson, et al., 2011; Pennycook et al., 2012; Pennycook et al., 2015). These results being consistent with previous base rate position and congruency results are important as it shows subjects did perform the base rate neglect task as instructed, and failure to find a difference between modalities was not due to a problem with the task itself or subjects’ performance in it.

**Base Rate Neglect Performance**

Subjects were more accurate in probability estimations when the base rate information was presented after the stereotype description than before. The present study’s results were consistent with several studies that also manipulated the ordering of base rate and stereotype description (Pennycook et al., 2015; Krosnick et al., 1990). When presented with base-rate
information prior to probability estimation, subjects are more likely to engage in a recognition of cognitive conflict, as the information used in conflict detection is more salient (Pennycook et al., 2015). However, our experiment did not find the significant main effect of base rate order for reaction, as seen in Pennycook et al., where in base rate last problems would have a faster reaction time for probability estimation. The inability to replicate this finding is due to the overall noise seen in the reaction time data, creating a large degree of variance in subject responses.

The current study also found a main effect for congruency where subjects were less accurate and took longer to respond when making probability estimations in incongruent trials than congruent trials. This is another typical finding in base rate neglect studies (e.g., Pennycook et al., 2015). In congruent trials there is no conflict of reasoning, as the base rate information is congruent to the personality description, leading to higher accuracy and faster reaction times. In incongruent trials, where there is a conflict between the base rate information and the stereotype description, subjects are more likely to experience a conflict in processing. This conflict engages Type 2 thinking, using cognitive decoupling to evaluate the base rate information and stereotype description separately and make a decision based on those two pieces of information.

**Experiment 2 Findings**

The primary objective of Experiment 2 was to attempt to increase the sensitivity to any performance differences in the base rate neglect task by increasing the amount of cognitive strain experienced during the task. I reasoned that if one of the modalities is more susceptible to working memory overload due to different working memory capacities, then any differences would be observable in the base rate neglect task. To this end, a memory load task was used to increase the amount of cognitive resource demands during the base rate neglect task. To ensure
and encourage subjects to perform the memory load task rather than cheat by ignoring it and devote their efforts solely on the base rate neglect task alone, an extra monetary incentive for reaching a certain accuracy threshold for performance on the memory task was added. The rest of Experiment 2 was the same as Experiment 1, except the base rate order manipulation was removed. The ordering for all base rate problems was base rate information, followed by the personality description. This was done to increase the sensitivity to the base rate neglect effect, by using the sequence of information that would most likely induce reasoning conflict. The results of the base rate neglect task were consistent with Experiment 1 and literature. Unexpectedly, there were modality differences in the memory load task. The following subsections will elaborate on those effects.

**Base Rate Neglect Performance**

The effects of congruency on task performance replicated the results of Experiment 1 and previous literature, where incongruent problems led to a decrease in accuracy and an increase in response time for probability estimations. In addition, there was a main effect of array size, where subjects were less accurate in probability estimations when the memory load task used a 3 letter array than a 7 letter array. Post-hoc analyses showed an interaction effect where base rate neglect task performance was only affected by array size in incongruent problems, to the effect of base rate task performance was worse when paired with memory load task arrays consisting of 3 items than arrays with 7 items. Congruent problems showed no difference in base rate performance between 3 and 7 item arrays. There was no statistical difference in base rate accuracy between modalities, which is unexpected given the rationale for Experiment 2. Instead, differences between modalities occurred in the secondary task.

**Memory Task Performance**
The purpose of the memory load task was to reduce the working memory capacity for performing the main base rate neglect task, and thus, impact estimate accuracy. To ensure subjects would perform the memory load task as accurately as possible, subjects were incentivized with extra money if their accuracy in the memory load task reached at least 81.25% (13 correct probes of the possible 16). Thus, due to the added performance incentive, I anticipated that subjects would prioritize the memory load task over the base rate neglect task, any deficits in working memory resources would more likely impact accuracy in the base rate neglect task. Moreover, I wanted to determine if the base rate performance differed between modalities when impaired through working memory load. This is not what I found. Similar to Experiment 1, Experiment 2 found no difference between modality groups in their accuracy of base rate estimations. Instead, I found a difference between modality groups in the memory load task where the auditory group was less accurate than the visual group, suggesting that memory load impacted only the auditory group. Overall, subjects were more accurate in the smaller array size condition than the larger array condition, as there were less items in the array that needed to be kept in short-term memory for the duration of the base rate neglect task. In addition, there was a main effect of congruency of the base rate neglect task on the memory load task performance, where when the base rate problem was congruent, subjects had lower memory load accuracy than when the base rate problem was incongruent. Post-hoc analyses revealed that this difference was due to the auditory modality group, where their performance on the memory load task was higher when paired with incongruent base rate problems. The visual group’s memory load performance was consistent with the expectations of performance between 3 item arrays and 7 items arrays and did not differ between congruency of the base rate neglect task.
Subjects in the visual modality were more accurate in the memory load task than subjects in the auditory modality. When looking at the post-hoc analyses, subjects in the visual modality across all conditions were highly accurate in the memory load task. There was a significant difference when looking at accuracy between array sizes; where there was a small decrease in accuracy in the larger array size condition compared to the smaller array condition. The subjects in the visual condition performed as expected, as their memory load task performance was unaffected by the base rate neglect task, whereas memory load task performance for the auditory group was influenced by the base rate neglect task.

To further assess performance in the memory task, its data was converted into Cowan’s K, a value representing the memory storage capacity in terms of number of items (Cowan, 2000). Results showed 7 item arrays had a higher memory storage capacity than 3 item arrays. While this might seem counterintuitive it actually makes sense given how Cowan’s K is calculated. The computed value K that estimates working memory capacity is limited by the number of items that are presented in the list. Congruency of the base rate neglect task also impacted the Cowan’s K value from the memory load task, where incongruent base rate problems led to a higher amount of memory capacity than congruent problems. Memory capacity was also different between modalities, where the visual group showed a higher degree of memory storage capacity than the auditory group. Moreover, the Array size by Modality by Congruency interaction reveals that the memory storage capacity of the visual modality is consistent across all experimental conditions and consistent with the expected results outlined. However, the poorer memory task performance of the auditory group in both accuracy and memory storage is unclear.

The visual group’s performance on the secondary task is indicative of previous literature regarding working memory differences between modalities, where visual modality presentation
requires less working memory demand than auditory presentations (Klingner et al., 2011). However, this was not the goal of the current study, as the purpose of the memory load task was to add additional working memory demands and impact performance on the base rate neglect task. Because the visual modality performed as expected on the memory load task, the unexpected results of the auditory modality cannot be due to the experiment design. This means that the difference in performance might be due to processing differences between modalities. One possible explanation may be the cross-modal demands of subjects in the auditory group switching between modalities when moving from the memory load task to the base rate neglect task. To explain, subjects in the auditory group were presented with visual stimuli in the memory load task and then with auditory stimuli in the base rate neglect task. This is unlike the visual group which had both tasks presented within the visual modality. It is possible that the cost to the memory load task performance in the auditory group was due to additional processing demands during cross-modality switching from visual stimulus processing to auditory stimulus processing. This explanation is consistent with other studies showing evidence of task performance costs when switching from one modality to another (Bragg & Redifer, 2022; Murray et al., 2009; Lukas et al., 2009; Sandhu & Dyson, 2012).

**Between Experiment Results**

Due to the lack of significant differences on base rate task performance between modalities in Experiment 1 and Experiment 2, analyses were conducted comparing base rate neglect performance between the experiments to determine if the memory load task had any effect on base rate neglect performance in Experiment 2 compared to Experiment 1. When comparing base rate accuracy, the data of Experiment 1 only included trials where the base rate information was presented first, to match the experiment design of Experiment 2. A comparison
of base rate accuracy between Experiment 1 and Experiment 2 showed that there was no difference in accuracy between experiments.

A comparison of the post-experimental survey between experiments was also conducted to determine if the subjects’ reported experience differed with the inclusion of the memory load task. Subjects reported greater overall task difficulty in Experiment 2 than Experiment 1. This is expected, as subjects in Experiment 2 are subjected to additional task demands with the inclusion of the memory load task in addition to the base rate neglect task. There was no difference in difficulty for remembering descriptions when making probability judgements between experiments. However, collapsing responses across experiments show a main effect for modality, where subjects in the auditory modality reported greater difficulty in remembering base rate problem description than subjects in the visual group.

Limitations and Future Directions

Even though the current study showed typical base rate neglect performance seen in previous base rate neglect literature, suggesting that the experimental design of the base rate neglect task was concrete, there are possible limitations within other parts of the experiment design which could be addressed in future studies. The overall context that this study took place in was abnormal, with having to make several concessions in controlling for other variables and limiting the methodological approaches I could use.

The overarching limitation that impacted the study design was conducting the experiment entirely online, due to the global pandemic of COVID-19. This was done for the safety of researchers and subjects alike, and in accordance with government and institutional guidelines in place for the duration of this study. Conducting the study online meant I was unable to include
certain methodologies and control for environmental factors during testing. In future studies, these measurements and considerations could be considered depending on restrictions in testing.

Because of the online nature of the study, it was not possible to fully control the environment subjects performed the task in. Even though it was recommended that subjects attend to the information on the screen and perform the task in an environment lacking distractions, it is not certain that testing took place under those conditions. In addition, subjects may have conducted the task with different monitor sizes or volume levels (in the auditory group) which could impact performance.

Had this study been conducted in the laboratory environment, the usage of eye tracking would have been possible. The addition of this measurement allows for tracking of subjects’ attention to the stimuli on the screen, both in the visual and auditory modalities. Experiment 1 did not find a significant main effect of base rate order for reaction time, as seen in Pennycook et al., where in base rate last problems would have a faster reaction time for probability estimation (2015). The inability to replicate this finding was due to the overall noise seen in the reaction time data, creating a large degree of variance in subject responses. Eye tracking and in lab testing would have provided some degree of security in ensuring that subjects were not distracted at the time of probability estimation. Not only would eye tracking would act as a measurement of attention, but also of cognitive load. As shown in previous literature, pupil dilation is an indicator of cognitive load in working memory (Klingner et al, 2011; Peavler, 1974; Kahneman & Beatty, 1966; for a review see van der Wel & Steenbergen, 2018). In the current study it was not observed on a behavioural level any differences between modality performance however, there still may have been different working memory load capacity usage between modalities; one method future research can use to capture that would be through pupil dilation.
Due to resource limitations, the inclusion of additional levels for several variables, which would have led to a more thorough experimental design, was not possible. One such inclusion would have been a dual modality presentation of the base rate problems, where problems are presented auditorily and visually. MacLeod et al. (2010) showed that verbally reciting materials while reading them led to better retention and recall of items studied. These results are due to the dual modal materials improving the encoding of information. This is due to a more salient encoding of the visual and auditory codes that would be associated with items, as compared to the visual and auditory representations encoded during solely visual stimuli with the phonological loop (MacLeod et al., 2010).

The unexpected performance of the auditory group in the memory load may have been due to possible cost demands associated with task-switching between modalities. Previous literature has shown that there is a cost in performance when task-switching is paired with a modality switch on a variety of different tasks (Bragg & Redifer, 2022; Murray et al., 2009; Lukas et al., 2009; Sandhu & Dyson, 2012). It may have been the case that subjects in the auditory group were more affected by the performance costs associated with task-switching. Some studies have shown that the auditory modality is more susceptible to performance costs associated with switching (Bragg & Redifer, 2022; Sandhu & Dyson, 2012; Lukas et al., 2009).

While not a specific limitation to the current study, there is still one task manipulation that could be included to increase future research’s sensitivity to cognitive load differences between modalities. Thompson et al., (2011) used a feeling of rightness measurement as a metacognitive monitoring measurement for detection conflict in the base rate neglect task. In future studies, a comparison of answer confidence between modalities may illuminate different capacities for conflict detection.
Conclusions

The present study’s primary focus was to determine if working memory differences exist in decision-making and reasoning processes between auditory and visual presentation modality, using a base rate neglect task. In both Experiment 1 and Experiment 2, differences in performance on the base rate neglect task were not found. However, Experiment 2, having the addition of a secondary task in the form of a memory load task, did capture a difference in performance where both the accuracy and memory storage capacity in the secondary task was lower for the auditory modality than the visual modality. A discussion of the results and limitations of the current study highlight how certain restrictions may have reduced the sensitivity of the experiment, and the capacity to detect differences in working memory during decision-making and reasoning processes. Future directions in this area of research should be done to consider other avenues of detecting working memory differences, and different task manipulations to place greater working memory demands on subjects to observe if working memory capacity differs between modalities during decision-making and reasoning.
References


Appendix A

Each base rate problem from Experiment 1.

1. Brannon is a randomly chosen participant of this study. Brannon is twenty nine years old. He is very good with numbers but is shy around people. He spends much of his time working. Among the participants of the study there were nine hundred ninety five street artists and five accountants.
   
   Base Rate Prompt A: Brannon is an accountant?
   
   Base Rate Prompt B: Brannon is a street artist?

2. Hank is a randomly chosen participant of this study. Hank is forty two years old. He is a creative and introverted person. He considers his home computer of his most prized possessions. Among the participants of the study there were nine hundred ninety six construction workers and four writers.

   Base Rate Prompt A: Hank is a writer?
   
   Base Rate Prompt B: Hank is a construction worker?

3. Tyrone is a randomly chosen participant of this study. Tyrone is twenty seven years old. All his friends consider him very brave and he is in relatively good physical shape. He goes to the gym regularly. Among the participants of the study there were nine hundred ninety six managers and four firemen.

   Base Rate Prompt A: Tyrone is a manager?
   
   Base Rate Prompt B: Tyrone is a fireman?

4. Geraldine is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven drummers and three librarians. Geraldine is forty one
years old. She loves books and spend a lot of her free time reading. She enjoys helping her two children with their homework.

Base Rate Prompt A: Geraldine is a drummer?
Base Rate Prompt B: Geraldine is a librarian?

5. Molly is a randomly chosen participants of this study. Among the participants of the study there were nine hundred ninety five researchers and five aerobics instructors. Molly is twenty five years old. She is very healthy and she works out at least five times a week. She enjoys pop music and dancing.

Base Rate Prompt A: Molly is a researcher?
Base Rate Prompt B: Molly is an aerobics instructor?

6. Jack is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five lawyers and five engineers. Jack is thirty six years old. He is not married and is somewhat introverted. He spends his free time reading science fiction and writing computer programs.

Base Rate Prompt A: Jack is a lawyer?
Base Rate Prompt B: Jack is an engineer?

7. Paul is a randomly chosen participant of this study. Among the participant of the study there were nine hundred ninety seven nurses and three doctors. Paul is thirty four years old. He lives in a beautiful home in a posh suburb. He is well spoken and very interested in politics.

Base Rate Prompt A: Paul is a doctor?
Base Rate Prompt B: Paul is a nurse?

8. Mike is a randomly chosen participant of this study. Mike is forty five years old. He is an imaginative person and enjoys street theatre. He loves experimenting with different types of
food and cooking. Among the participants of the study there were nine hundred ninety seven consultants and three artists.

Base Rate Prompt A: Mike is an artist?
Base Rate Prompt B: Mike is a consultant?

9. George is a randomly chosen participant of this study. George is thirty six years old. He is very intelligent and has nerves of steel. He has great hand eye coordination and spatial awareness. Among the participant of the study there were nine hundred ninety six aeroplane pilots and four shop assistants.

Base Rate Prompt A: George is a shop assistant?
Base Rate Prompt B: George is an aeroplane pilot?

10. Lilly is a randomly chosen participant of this study. Lilly is thirty seven years old. She is married and has several kids. Her husband is a veterinarian. She always watches cartoons with her kids. Among the participants of the study there were nine hundred ninety six kindergarten teachers and four executive managers.

Base Rate Prompt A: Lilly is a kindergarten teacher?
Base Rate Prompt B: Lilly is an executive manager?

11. Richard is a randomly chosen participant of this study. Richard is fifty six years old. He is a good public speaker and is good at meeting people. He can argue both sides of an issue. Among the participants of the study there were nine hundred ninety five politicians and five I. T. technicians.

Base Rate Prompt A: Richard is a politician?
Base Rate Prompt B: Richard is an I. T. technician?
12. Lucius is a randomly chosen participant of this study. Lucius is thirty four years old. He is prone to getting injuries. He recently got divorced. He goes for a run every morning to exercise. Among the participants of the study there were nine hundred ninety seven boxers and three hippies.

   Base Rate Prompt A: Lucius is a boxer?
   Base Rate Prompt B: Lucius is a hippy?

13. Dan is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five paramedics and five clowns. Dan is thirty two years old. He is a good driver and takes his job very seriously. He is married but has no children.

   Base Rate Prompt A: Dan is a clown?
   Base Rate Prompt B: Dan is a paramedic?

14. Martine is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five French people and five Americans. Martine is twenty six years old. She is bilingual and reads in her spare time. She is a very fashionable dresser and a great cook.

   Base Rate Prompt A: Martine is French?
   Base Rate Prompt B: Martine is American?

15. Corinne is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety six secretaries and four gardeners. Corinne is thirty two years old. She is a great organizer and always dresses neatly. She loves talking to her friends and family on the phone.

   Base Rate Prompt A: Corinne is a secretary?
   Base Rate Prompt B: Corinne is a gardener?
16. Dianna is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven nannies and three telemarketers. Dianna is sixty years old. She loves children and has had her current job for several years. She enjoys drinking tea and visiting with family and friends.

Base Rate Prompt A: Dianna is a telemarketer?

Base Rate Prompt B: Dianna is a nanny?

Practice Trial A. Kelly is a randomly chosen participant of this study. Among the participants of the study there were nine hundred and ninety seven girls and three boys. Kelly is thirteen years old. Kelly’s favourite subject is art. Kelly’s favourite activities are shopping and having sleepovers with friends to gossip about other kids at school.

Base Rate Prompt: Kelly is a girl?

Practice Trial B. Jessie is a randomly chosen participant of this study. Jessie is twenty three years old and is finishing a degree in engineering. Jessie likes to go out with friends listen to loud music and drink beer. Among the participants of the study there were nine hundred ninety six women and four men.

Base Rate Prompt: Jessie is a man?
Appendix B

Experiment 1 Instructions.

In this experiment you will be presented with a series of descriptions about an individual, their hobbies and personality, and a study they participated in, which consisted of two population groups (e.g., mechanics and musicians).

After the description is presented you will be asked to judge the probability that the person mentioned is of one of the two groups mentioned in the description. This judgement will be as a percentage. For example, a judgement of "0" means that the probability is "0%", whereas "100" means "100%".

Auditory Modality- Each description will be presented auditorily by a male voice.

Visual Modality- Each description will be presented in a rapid serial presentation format, where each word will be presented on the screen one at a time.

This is the slider that you will use to judge the probability of the person being part of one of the two specific groups mentioned, as a percentage.

Click on the slider for the marker to appear. You can then click and drag it anywhere along the scale. Once you have made your selection, press the "space bar" to continue to the next description.

The next screen will be two practice trials to get you familiar with the experiment.

The next screen will be the start of the experiment. It will consist of 16 trials in total, and takes approximately 10 minutes to complete.
Appendix C

Each base rate problem from Experiment 2 and the memory load prompt and probe associated with each base rate problem.

1. Brannon is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five street artists and five accountants. Brannon is twenty nine years old. He is very good with numbers but is shy around people. He spends much of his time working.

   Base Rate Prompt A: Brannon is an accountant?
   Base Rate Prompt B: Brannon is a street artist?
   Memory Load: JWH
   Memory Load Probe: S

2. Hank is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety six construction workers and four writers. Hank is forty two years old. He is a creative and introverted person. He considers his home computer of his most prized possessions.

   Base Rate Prompt A: Hank is a writer?
   Base Rate Prompt B: Hank is a construction worker?
   Memory Load: BVP
   Memory Load Probe: R

3. Tyrone is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety six managers and four firemen. Tyrone is twenty seven years old. All his friends consider him very brave and he is in relatively good physical shape. He goes to the gym regularly.
4. Geraldine is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven drummers and three librarians. Geraldine is forty one years old. She loves books and spends a lot of her free time reading. She enjoys helping her two children with their homework.

   Base Rate Prompt A: Geraldine is a drummer?
   Base Rate Prompt B: Geraldine is a librarian?
   Memory Load: MPQ
   Memory Load Probe: M

5. Molly is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five researchers and five aerobics instructors. Molly is twenty five years old. She is very healthy and she works out at least five times a week. She enjoys pop music and dancing.

   Base Rate Prompt A: Molly is a researcher?
   Base Rate Prompt B: Molly is an aerobics instructor?
   Memory Load: WXUCTXG
   Memory Load Probe: S

6. Jack is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five lawyers and five engineers. Jack is thirty six years old.
He is not married and is somewhat introverted. He spends his free time reading science fiction and writing computer programs.

Base Rate Prompt A: Jack is a lawyer?
Base Rate Prompt B: Jack is an engineer?
Memory Load: RDL
Memory Load Probe: L

7. Paul is a randomly chosen participant of this study. Among the participant of the study there were nine hundred ninety seven nurses and three doctors. Paul is thirty four years old. He lives in a beautiful home in a posh suburb. He is well spoken and very interested in politics.

Base Rate Prompt A: Paul is a doctor?
Base Rate Prompt B: Paul is a nurse?
Memory Load: OKBLFGJ
Memory Load Probe: N

8. Mike is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven consultants and three artists. Mike is forty five years old. He is an imaginative person and enjoys street theatre. He loves experimenting with different types of food and cooking.

Base Rate Prompt A: Mike is an artist?
Base Rate Prompt B: Mike is a consultant?
Memory Load: LOZMQKA
Memory Load Probe: K

9. George is a randomly chosen participant of this study. Among the participant of the study there were nine hundred ninety six aeroplane pilots and four shop assistants. George is
thirty six years old. He is very intelligent and has nerves of steel. He has great hand eye
coordination and spatial awareness.

Base Rate Prompt A: George is a shop assistant?
Base Rate Prompt B: George is an aeroplane pilot?
Memory Load: MTY
Memory Load Probe: T

10. Lilly is a randomly chosen participant of this study. Among the participants of the
study there were nine hundred ninety six kindergarten teachers and four executive managers.
Lilly is thirty seven years old. She is married and has several kids. Her husband is a veterinarian.
She always watches cartoons with her kids.

Base Rate Prompt A: Lilly is a kindergarten teacher?
Base Rate Prompt B: Lilly is an executive manager?
Memory Load: AQJ
Memory Load Probe: A

11. Richard is a randomly chosen participant of this study. Among the participants of the
study there were nine hundred ninety five politicians and five I. T. technicians. Richard is fifty
six years old. He is a good public speaker and is good at meeting people. He can argue both sides
of an issue.

Base Rate Prompt A: Richard is a politician?
Base Rate Prompt B: Richard is an I. T. technician?
Memory Load: JHI
Memory Load Probe: X
12. Lucius is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven boxers and three hippies. Lucius is thirty four years old. He is prone to getting injuries. He recently got divorced. He goes for a run every morning to exercise.

Base Rate Prompt A: Lucius is a boxer?
Base Rate Prompt B: Lucius is a hippy?
Memory Load: CVO
Memory Load Probe: W

13. Dan is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five paramedics and five clowns. Dan is thirty two years old. He is a good driver and takes his job very seriously. He is married but has no children.

Base Rate Prompt A: Dan is a clown?
Base Rate Prompt B: Dan is a paramedic?
Memory Load: TPKGBDK
Memory Load Probe: P

14. Martine is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety five French people and five Americans. Martine is twenty six years old. She is bilingual and reads in her spare time. She is a very fashionable dresser and a great cook.

Base Rate Prompt A: Martine is French?
Base Rate Prompt B: Martine is American?
Memory Load: JLARGNT
Memory Load Probe: G
15. Corinne is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety six secretaries and four gardeners. Corinne is thirty two years old. She is a great organizer and always dresses neatly. She loves talking to her friends and family on the phone.

   Base Rate Prompt A: Corinne is a secretary?
   Base Rate Prompt B: Corinne is a gardener?
   Memory Load: LCKMHXI
   Memory Load Probe: E

16. Dianna is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety seven nannies and three telemarketers. Dianna is sixty years old. She loves children and has had her current job for several years. She enjoys drinking tea and visiting with family and friends.

   Base Rate Prompt A: Dianna is a telemarketer?
   Base Rate Prompt B: Dianna is a nanny?
   Memory Load: NOJTSBC
   Memory Load Probe: H

Practice Trial A. Kelly is a randomly chosen participant of this student. Among the participants of the study there were nine hundred and ninety seven girls and three boys. Kelly is thirteen years old. Kelly’s favourite subject is art. Kelly’s favourite activities are shopping and having sleepovers with friends to gossip about other kids at school.

   Base Rate Prompt: Kelly is a girl?
   Memory Load: KCFPEJX
   Memory Load Probe: C
Practice Trial B. Jessie is a randomly chosen participant of this study. Among the participants of the study there were nine hundred ninety six women and four men. Jessie is twenty three years old and is finishing a degree in engineering. Jessie likes to go out with friends listen to loud music and drink beer.

Base Rate Prompt: Jessie is a man?

Memory Load: DKF

Memory Load Probe: O
Appendix D

Experiment 2 Instructions.

In this experiment you will be presented with a series of descriptions about an individual, their hobbies and personality, and a study they participated in, which consisted of two population groups (e.g., mechanics and musicians).

After the description is presented you will be asked to judge the probability that the person mentioned is of one of the two groups mentioned in the description. This judgement will be as a percentage. For example, a judgement of "0" means that the probability is "0%", whereas "100" means "100%".

Auditory Modality- Each description will be presented auditorily by a male voice.

The next screen will be the voice that you will hear for each trial. Please make sure your speakers are turned on.

Visual Modality- Each description will be presented in a rapid serial presentation format, where each word will be presented on the screen one at a time.

This is the slider that you will use to judge the probability of the person being part of one of the two specific groups mentioned, as a percentage.

Click on the slider for the marker to appear. You can then click and drag it anywhere along the scale. Once you have made your selection, press the "space bar" to continue to the next description.

At the beginning of each trial, you will also be shown a string of letters for three seconds. These letter strings will be either three or seven letters long. Here is an example:

ABCDEFG
After giving your probability judgement you will be shown a single letter. You must respond with the '1' key if the shown letter was in letter string for that trial, and the '2' key if the letter was not shown in the letter string for that trial. Here is an example:

Was the letter 'D' shown?
Press '1' for yes, or '2' for no.

The next screen will be two practice trials to get you familiar with the experiment.

It is very important for you to respond as accurately as possible to the questions about the letter prompt in addition to paying attention to the description.

To add an incentive for accurate performance, participants who get 80% or better in the letter prompts will receive a bonus reward of 1€ in addition to the regular payment.

The next screen will be the start of the experiment. It will consist of 16 trials in total, and takes approximately 10 minutes to complete.