

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND
INFORMATICS

THE ROLE OF PRIOR BELIEFS ON MEMORY RECALL

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND
INFORMATICS

THE ROLE OF PRIOR BELIEFS ON MEMORY RECALL
DIPLOMA THESIS

Study program: Cognitive Science

Field of study: Computer Science

Supervising Department: Department of Applied Informatics

Supervisor: Prof. Christophe Heintz, Ph.D.

Consultant: Francesca Bonalumi, Ph.D.

Bratislava 2024

Kutlay Usta



Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Bc. Kutlay Usta
Study programme: Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study: Computer Science
Type of Thesis: Diploma Thesis
Language of Thesis: English
Secondary language: Slovak

Title: The Role of Prior Beliefs on Memory Recall

Annotation: This thesis investigates how prior beliefs, particularly the belief in a just world, affect memory recall and false memory formation. It examines if individuals with strong beliefs in a fair world create false memories when these beliefs are challenged. The study also explores how time delays and commitment fulfillment affect memory reconstruction, using an algorithm to assess recall accuracy and implicature.

Aim:

1. Gain a deeper understanding of the influence of prior beliefs on memory.
2. Explore the effect of commitments and time gaps on memory recall.
3. Develop a new assessment method for memory recalls that assists in minimizing human errors.

Literature:

Roediger, H. L., & McDermott, K. B. (1996). Creating false memories: Remembering words not presented in lists.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories.

Festinger, L. (1957). A theory of cognitive dissonance.

Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition (2nd ed.).

Supervisor: Prof. Christophe Heintz, Dr.
Consultant: Francesca Bonalumi, PhD.
Department: FMFI.KAI - Department of Applied Informatics
Head of department: Prof. Ing. Igor Farkaš, Dr.
Assigned: 23.05.2024

Approved: prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

.....
Student

.....
Supervisor



ZADANIE DIPLOMOVEJ PRÁCE

Meno a priezvisko:	Bc. Kutlay Usta
Študijný program:	(Jednoodborové štúdium, magisterský II. stupeň, plný časová forma)
Študijný odbor:	Informatika
Typ diplomovej práce:	Diplomová práca
Jazyk práce:	Jazyk práce: angličtina
Druhý jazyk:	Slovenský jazyk
Názov:	Úloha predchádzajúcich presvedčení pri vybavovaní pamäte
Anotácia:	Táto práca skúma, ako predchádzajúce presvedčenia, najmä presvedčenie o spravodlivom svete, ovplyvňujú vybavovanie si spomienok a tvorbu falošných spomienok. Skúma, či jednotlivci so silným presvedčením o spravodlivom svete vytvárajú falošné spomienky, keď sú tieto presvedčenia spochybnené. Štúdia tiež skúma, ako časové oneskorenie a plnenie záväzkov ovplyvňuje rekonštrukciu pamäti, pričom používa algoritmus na hodnotenie presnosti vybavovania a implikácie.
Ciel:	<ol style="list-style-type: none">1. Hlbšie pochopiť vplyv predchádzajúcich presvedčení na pamäť.2. Preskúmajte vplyv záväzkov a časových medzier na zapamätanie si.3. Vyvinúť novú metódu hodnotenia pamäťových spomienok, ktorá pomáha priminimalizácia ľudských chýb.
Literatúra:	<p>Roediger, H. L., & McDermott, K. B. (1996). Vytváranie falošných spomienok: Zapamätanie si slov, ktoré nie sú uvedené v zoznamoch.</p> <p>Loftus, E. F., & Pickrell, J. E. (1995). Tvorba falošných spomienok.</p> <p>Festinger, L. (1957). Teória kognitívnej disonancie.</p> <p>Sperber, D., & Wilson, D. (1995). Relevantnosť: (2. vyd.).</p>
Dozor:	Christophe Heintz, Dr.
Konzultant:	Francesca Bonalumi, PhD.
Oddelenie:	FMFI.KAI - Katedra aplikovanej informatiky
Vedúci oddelenia:	Prof. Ing. Igor Farkaš, Dr.
Pridelené:	23.05.2024
Schválené:	Ing. Igor Farkaš, Dr. Garant študijného programu

..... Študent Supervízor

Študent

Supervízor

Declaration

I have written and investigated this thesis using relevant literature that is adequately cited and quoted where needed. Additionally, I used Grammarly AI (grammarly.com) during the writing process to improve and fix the grammar, engage the sentences, and ensure they sound fluent.

Bratislava, 2024

Kutlay Usta

Acknowledgments

I would like to thank my supervisors, Professor Christophe Heintz and Dr. Francesca Bonalumi, for their invaluable support throughout the process of writing my thesis. I also thank the Adaptive Cognition & Economics in Society (ACES) members for their valuable assistance. Lastly, I would like to acknowledge the support of my family, particularly Necla Selat, who recently passed away, for their support during this process.

Abstract

In this study, we aimed to investigate the role of prior beliefs on memory recall and the reconstructive processes involved, mainly focusing on how a prior belief in a just world and how commitment violations affect false memory formation during reconstructive memory recall. We hypothesized that individuals who strongly believe in a fair world would be motivated to create false memories (and recall an implicature being explicitly uttered) when their beliefs are violated. Additionally, we hypothesized that individuals would similarly reconstruct their memories if commitments were not satisfied and there was a time delay between exposure and recall. To test these hypotheses, we developed an algorithm that automatically evaluates the discrepancies between recalled and target sentences regarding accuracy and implicature, thereby reducing the possibility of human error in assessment. The algorithm checks for grammatical distinctions to measure accuracy and examines differences in meaning arising from pragmatic inferences to measure implicature. Participants were presented with vignettes and asked to recall the target sentences under broken, kept, delay, and no-delay conditions. Our results did not support the hypothesis regarding the belief in a just world and false memories. However, they highlighted the significant roles of time delay and the nature of commitments in shaping memory recall. Participants with a time delay showed higher implicature scores and lower accuracy, indicating that time influences the reconstruction of memories towards more pragmatic inferences. Interestingly, participants recalled broken commitments more accurately than kept ones, likely due to the heightened emotional and cognitive impact of violated expectations. These findings underscore the complexity of memory processes and the need for further research to disentangle the various cognitive and motivational factors involved. Our development of an automatic assessment method for memory recall offers a valuable tool for future studies, enhancing the reliability and reproducibility of research in this field.

Keywords: memory recall, false memories, just world belief, implicature, motivated memory

Abstrakt

V tejto štúdii sme sa zamerali na skúmanie úlohy predchádzajúcich presvedčení pri vybavovaní si spomienok a súvisiacich rekonštrukčných procesov, pričom sme sa zamerali najmä na to, ako predchádzajúce presvedčenie o spravodlivom svete a porušenie záväzkov ovplyvňujú tvorbu falošnej pamäte počas rekonštrukčného vybavovania si spomienok. Predpokladali sme, že jednotlivci, ktorí pevne veria v spravodlivý svet, budú motivovaní vytvárať falošné spomienky (a vybavovať si implikáciu, ktorá bola explicitne vyslovená), keď sú ich presvedčenia porušené. Okrem toho sme predpokladali, že jednotlivci budú podobne rekonštruovať svoje spomienky, ak záväzky nebudú splnené a medzi vystavením a vyvolaním bude časový odstup. Na overenie týchto hypotéz sme vyvinuli algoritmus, ktorý automaticky vyhodnocuje nezrovnalosti medzi vyvolanými a cieľovými vetami, pokiaľ ide o presnosť a implikáciu, čím sa znižuje možnosť ľudskej chyby pri vyhodnocovaní. Algoritmus kontroluje gramatické rozdiely na meranie presnosti a skúma rozdiely vo význame vyplývajúce z pragmatických záverov na meranie implikatúry. Účastníkom boli predložené vinety a boli požiadani, aby si spomenuli na cieľové vety v podmienkach s preruseným, zachovaným, oneskoreným a žiadnym oneskorením. Naše výsledky nepotvrdili hypotézu týkajúcu sa viery v spravodlivý svet a falošných spomienok. Zdôraznili však významnú úlohu časového oneskorenia a povahy záväzkov pri formovaní spomienkového vybavovania. Účastníci s časovým oneskorením vykazovali vyššie skóre implikácií a nižšiu presnosť, čo naznačuje, že čas ovplyvňuje rekonštrukciu spomienok smerom k pragmatickejším záverom. Zaujímavé je, že účastníci si porušené záväzky vybavovali presnejšie ako dodržané, pravdepodobne v dôsledku zvýšeného emocionálneho a kognitívneho vplyvu porušených očakávaní. Tieto zistenia zdôrazňujú zložitosť pamäťových procesov a potrebu ďalšieho výskumu na rozlíšenie rôznych kognitívnych a motivačných faktorov, ktoré sa na nich podielajú. Nás vývoj metódy automatického hodnotenia pamäťového vybavovania ponúka cenný nástroj pre budúce štúdie, ktorý zvyšuje spoľahlivosť a reprodukovateľnosť výskumu v tejto oblasti.

Kľúčové slová: vybavovanie si spomienok, falošné spomienky, presvedčenie o spravodlivom svete, implikatúra, motivovaná pamäť.

Forewords

This thesis aims to understand the sophisticated processes of memory formation and recall. It focuses on the belief in a just world, commitment violations, and their impact on how individuals reconstruct memories of events. Specifically, it explores whether such beliefs create false memories when expectations are violated. The thesis also investigates the roles of time delay and the nature of commitments (broken or kept) in shaping memory accuracy and implicature formation.

As an innovative aspect, using an automatic assessment method to evaluate memory recall enhances the reliability and reproducibility of the findings. By integrating cognitive and motivational perspectives, this study offers valuable insights into the reconstructive nature of memory and provides a novel methodological approach for future research.

We hope this thesis will contribute to the growing body of literature on memory studies and inspire further exploration of cognitive and motivational factors in memory processes.

Contents

INTRODUCTION	1
1 LITERATURE REVIEW.....	2
1.1 THEORIES OF MEMORY	2
1.1.1 The Stages of Memory: Encoding, Storage, and Retrieval.....	2
1.1.2 Schema Theory.....	6
1.1.3 Dual-Processing Model	7
1.1.4 Fuzzy-trace Theory.....	8
1.1.5 Self-Memory System Model	9
1.1.6 The Source-Monitoring Framework.....	9
1.2 FALSE MEMORIES	10
1.2.1 Type of Memory Malfunctions.....	11
1.2.2 Schema Theory.....	13
1.2.3 The Source-Monitoring Framework.....	16
1.2.4 Dual-Process Model	17
1.2.5 Fuzzy-trace Theory.....	18
1.2.6 Self-Memory System Model	20
1.2.7 Neurobiological Perspectives of False Memories	22
1.3 MOTIVATED MEMORY	24
1.3.1 Motivated Forgetting and Misremembering	24
1.3.2 Cognitive Dissonance.....	26
1.3.3 Belief in a Just World	27
1.3.4 Accountability	32
1.4 LINGUISTICS AND MEMORY	34
1.4.1 Semantics and Pragmatics	35
1.4.2 Implicature	35
1.4.3 Influence of Time	36
2 CURRENT STUDY	38
3 METHOD	41
3.1 PARTICIPANTS	41
3.2 MATERIALS	41
3.2.1 Vignettes.....	41

3.2.2 Memory Recall Question.....	44
3.2.3 Attention Check Question	46
3.2.4 Just World Scale	46
3.2.5 Implicature Scoring Algorithm.....	46
3.3 EXPERIMENTAL DESIGN	51
3.3.1 No Delay Conditions	51
3.3.2 Delay Conditions.....	52
4 RESULTS.....	53
4.1 JUST WORLD HYPOTHESIS	54
4.2 MOTIVATED-BY-ACCOUNTABILITY-GOALS HYPOTHESIS	55
4.3 MEANING-TO-SENTENCE HYPOTHESIS	56
4.4 OTHER REMARKS.....	58
5 DISCUSSION.....	61
5.1 FORGETTING	61
5.2 IMPLICATURE SCORING ALGORITHM.....	61
5.3 JUST WORLD HYPOTHESIS	62
5.4 MOTIVATED-BY-ACCOUNTABILITY-GOALS HYPOTHESIS	65
5.5 MEANING-TO-SENTENCE HYPOTHESIS	66
5.5 LIMITATIONS	67
CONCLUSION	70
REFERENCES.....	72
APPENDIX A - VIGNETTES	84
APPENDIX B – ETHICAL APPROVAL.....	92
APPENDIX C – ATTENTION CHECK QUESTIONS	93
APPENDIX D – CONSENT FORM	95
APPENDIX E – GITHUB REPOSITORY	96
APPENDIX F – BELIEF IN A JUST WORLD SCALE	97
APPENDIX G – PARTICIPANT DISTRIBUTION.....	99
APPENDIX H – DETAILED ANALYSIS OF VIGNETTES	100

List of Tables

TABLE 1 THE NORMALITY TEST OF COMMITMENT AND DELAY CONDITIONS	54
---	----

List of Figures

FIGURE 1 THE GRAPHICAL REPRESENTATION OF JUST WORLD HYPOTHESIS	39
FIGURE 2 FIRST NORMING STUDY THREE-POINT QUESTIONNAIRE ANSWER DISTRIBUTION	44
FIGURE 3 ACCURACY SCORING PROCESS FLOWCHART.....	48
FIGURE 4 IMPLICATURE SCORING PROCESS FLOWCHART	49
FIGURE 5 THE FLOWCHART OF EXPERIMENT PROCEDURE	52
FIGURE 6 THE TREE OF CHI-SQUARE TESTS	53
FIGURE 7 QUANTILE REGRESSION FOR IMPLICATURE SCORES IN BROKEN COMMITMENT	55
FIGURE 8 MANN-WHITNEY U TEST FOR IMPLICATURE SCORES BY COMMITMENT CONDITION	56
FIGURE 9 MANN-WHITNEY U TEST FOR IMPLICATURE SCORES BY DELAY CONDITION	57
FIGURE 10 MANN-WHITNEY U TEST FOR ACCURACY SCORES BY DELAY CONDITION	57
FIGURE 11 QUANTILE REGRESSION FOR IMPLICATURE SCORES IN KEPT COMMITMENT	58
FIGURE 12 DECISION TREE OF DELAY AND COMMITMENT CONDITIONS ON IMPLICATURE.....	59
FIGURE 13 QUANTILE REGRESSION FOR ACCURACY SCORES IN BROKEN COMMITMENT	59
FIGURE 14 QUANTILE REGRESSION FOR ACCURACY SCORES IN KEPT COMMITMENT	60
FIGURE 15 MANN-WHITNEY U TEST FOR ACCURACY SCORES BY COMMITMENT CONDITION.....	60
FIGURE 16 COMPARISON OF MEANS OF IMPLICATURE AND ACCURACY SCORES.....	102
FIGURE 17 THE PERCENTAGE OF THE TOTAL COUNT OF ERROR FOR EACH VIGNETTE	103

Abbreviations And Symbols

ACC - Anterior Cingulate Cortex

aINS - Anterior Insula

BJW - Belief in a Just World

BJW_D - Belief in a Just World Distributive Justice

BJW_P - Belief in a Just World Procedural Justice

BJW-others - Belief in a Just World for Others

BJW-self - Belief in a Just World for Self

IMS – Implicature Scoring Algorithm

M - Mean

N - Number of Participants

p - p-value (probability value in statistical hypothesis testing)

PCC - Posterior Cingulate Cortex

PTSD - Post-Traumatic Stress Disorder

SD - Standard Deviation

TOT - Tip-of-the-Tongue

α - Cronbach's Alpha (a measure of reliability or internal consistency)

χ^2 - Chi-Square Test

Introduction

How do people remember events? What processes are involved in forming memories? Do prior beliefs influence these processes? How can preexisting knowledge, motivations, and other cognitive factors affect memory recall? In this thesis, we will explore those questions and their answers. We will examine the literature on memory, memory malfunctions, false memory formation, and their cognitive perspectives, followed by the motivational and linguistic influence on memory. Lastly, we will present our current study, hypotheses, and findings.

We hypothesized that sentence recall is affected by a memory reconstruction process driven by pre-existing beliefs that the world is fair (Lerner, 1980). To test our hypothesis, we designed an experiment where participants were instructed to recall a specific sentence uttered by a protagonist in a short story that they read. Notably, the sentence implied that the protagonist committed to do something. We manipulated the story's outcomes regarding broken or kept commitment, assessing the influence of their pre-existing beliefs on the recall.

As a novel approach, we developed an automatic assessment method that assesses the relations between a sentence and what people remember in two aspects: Accuracy and Implicature. The accuracy score indicates grammatical differences between commitment sentences and participant recall, and the Implicature score indicates the extent of pragmatic inferences.

With this experiment, we aimed to understand how people's beliefs about the world impact the formation of false memories in reconstructive aspects of memory recall. Moreover, by exploring this topic and developing an assessment method, we aim to contribute to the memory literature with our findings and a more effective assessment method to help further studies.

1 Literature Review

1.1 Theories of Memory

Memory is not a single monolithic system but a complex interplay of various processes and stages. The stages of memory—encoding, storage, and retrieval—each play a distinct role in how the brain handles information (Loftus et al., 1976). Various cognitive theories and models have been developed to explain these processes, including the stage model, schema theory, dual-processing model, fuzzy-trace theory, the self-memory system, and the source-monitoring framework. Each model offers a unique perspective on how memories are formed, maintained, and sometimes distorted. In this section, we will explore the theories and models that explain the mechanisms of memory before delving into their explanation of the formation of false memories.

1.1.1 The Stages of Memory: Encoding, Storage, and Retrieval

Encoding is the first process of memory, where information is initially collected from sensory inputs and transformed into a form that can be stored in memory. However, not every piece of information is encoded. Instead, an individual's attention, motivation, prior knowledge, and experiences impact the encoding process. More specifically, how we attend to and interpret information influences how we encode it in our memory (Roediger et al., 1996; Stangor & Walinga, 2014).

Storing is the second stage of the memory process, referring to the information-retaining process (Roediger et al., 1996). Following the encoding of sensory inputs, the brain selectively encodes the information and transfers it to short-term memory. Subsequently, relevant information is transferred to long-term memory, where they are stored for extended periods and can be easily retrieved.

The retrieval is the third and final stage of the memory process, which refers to accessing stored information from long-term or short-term memory, allowing the individual to acquire tasks or purposes (Atkinson & Shiffrin, 1968).

Various factors influence the retrieval process, such as the strength and clarity of the memory trace, the context in which the information was learned, and the individual's motivation and attention. Additionally, interference, forgetting, and memory distortion affect the quality of the retrieved information (Craik & Lockhart, 1972). The retrieval process can occur either voluntarily or involuntarily. Voluntary memory retrieval occurs when an individual tries to remember a piece of information, such as a name. Involuntary memory retrieval occurs when an individual unexpectedly recalls a memory with the influence of a trigger, such as a smell, sound, or sight.

Atkinson and Shiffrin (1968) propose that memory is a multi-stage process that involves manipulating novel information before retention. Therefore, the stage model suggests three stages of memory: sensory, short-term, and long-term memory.

Interactions between cognition and the environment initially start through sense organs. The sensory organs, such as eyes or ears, collect auditory and visual information about the outer world. Memory unselectively stores all information perceived through sensory organs for a brief period, usually lasting within a second. This temporary information storage process constitutes the first stage of the memory system, called sensory memory (Loftus et al., 2019). Some information stored in sensory memory transfers to the short-term memory before fading. Attention and motivation are distinctive factors in whether information fades or transfers to short-term memory.

For instance, when we look up our names in a name list, the information is perceived through sensory organs such as our eyes. In such a case, the information on our names and other information, such as different names in the list, are entered into the sensory memory. However, the collected information about other names decays away unless transferred to short-term memory, usually within a second.

Short-term memory is a stage in which limited information is fed by sensory memory. It is a temporary respiratory stage in which most information typically decays less than a minute before being transferred to long-term memory (Atkinson & Shiffrin, 1968; Baddeley, 2006; Lutz & Huitt, 2003). For instance, individuals often repeat new names when they search for them in a name list to avoid potential memory lapses.

On the other hand, working memory is a cognitive system that performs various tasks with the information temporarily stored in short-term memory. It helps individuals to retrieve information from long-term memory and transfer it back (Baddeley, 2006). The duration of information storage in short-term memory affects the transfer process (Atkinson & Shiffrin, 1968).

Baddeley and Hitch (2007) divided working memory into three components regarding their functions: the visuospatial sketchpad, central executive, and phonological loop.

The phonological loop is essential in various cognitive processes such as language learning, reading, spelling, arithmetic, and reasoning. It is responsible for processing auditory and verbal information, such as words, numbers, and sounds. The phonological loop comprises two subcomponents: a phonological store that temporarily retains auditory or verbal information and an articulatory rehearsal process that maintains the information in the phonological store through repeating words or phrases in mind. For instance, when we hear a vital phone number, the phonological loop helps us retain the information in our memory as we prepare to call it.

The visuospatial sketchpad is a part of working memory that stores and manipulates visual and spatial information. For instance, when we try to visualize a map to find our way to a new location, the visuospatial sketchpad helps us see various processes, such as images of the map, shapes of the buildings, colors of the surroundings, locations of the landmarks, distances between them, and movements required to reach our destination.

The central executive part is the most crucial component of working memory that controls and coordinates the phonological loop and visuospatial sketchpad. It retrieves and manipulates information only a few items at a time while interacting with long-term memory. However, the central executive does not function as a solid system. Instead, a collection of functions performs various tasks such as attention, inhibition, planning, and decision-making.

The long-term memory stage is the last stage of the memory system. The characteristic difference between the rest of the stages is the duration of information in this memory. As we discussed, most information temporarily stored in sensory and short-term memory eventually decays. In contrast, the information stored in long-term memory is relatively permanent and may range from a few minutes to a lifetime (Atkinson & Shiffrin, 1968). The other stages feed

the stage of long-term memory. For instance, before we buy an apple, we smell it (sensory memory), check the shape (sensory, short-term, and working memory), and compare it with the smell and shape of the apples we ate a month ago (long-term memory).

Long-term memory is responsible for various memory tasks in cognition. Regarding their characteristics, Squire et al. (1993) divided these memory tasks into two subcomponents of long-term memory: explicit and implicit.

Explicit memory (declarative memory) is a component of long-term memory that involves conscious recollection of events, facts, and experiences (Schacter, 1989). As a characteristic, explicit memory is slowly accessible and has limited capacity (Colombo et al., 2000). It is mainly associated with remembering (Gardiner & Java, 1993). Consequently, it is often studied using specific memory tasks such as recognition and recall, which involve remembering information about something that is not currently perceived, such as retrieving a phone number or identifying a person's name from a list of names (Mandler, 1990; Colombo et al., 2000).

According to Tulving (1972), explicit memory has two subsystems: episodic and semantic. Episodic memory is a memory responsible for autobiographical events. It provides contextual information about an individual's past experiences, such as feelings and thoughts (Braisby et al., 2012). Beyond that, episodic memory allows individuals to re-experience their past experiences and mental time travel (Baddeley et al., 2009).

Furthermore, recent theories suggest that episodic memory also serves a communicative function. According to Mahr and Csibra (2018), episodic memory is about recalling past events and justifying our beliefs and assertions about these events in social contexts. This communicative account posits that episodic memory helps individuals establish epistemic authority and accountability in their interactions. It provides detailed and self-referential recollections in which individuals can demonstrate their memories' reliability. Thus, they can strengthen their trust and cooperation within social groups. This perspective aligns with the broader role of memory in supporting social functions, such as holding others accountable for their actions (Cosmides & Tooby, 1992). In this view, episodic memory's ability to construct and communicate detailed narratives of past experiences is crucial for maintaining social bonds and ensuring the integrity of social exchanges.

Semantic memory contains an organized knowledge of words and other verbal symbols, such as their meanings, relationships, and rules for manipulating them. It allows individuals to retrieve information that is not directly stored without changing its contents. The semantic system is more resistant to involuntary alteration and loss of information than the episodic system, and it functions independently of the episodic system to record and maintain information (Tulving, 1972).

For instance, "Have you ever been to Bratislava?" challenges the ability to recall a specific experience, a function of episodic memory. On the other hand, "What is the capital of Slovakia?" tests the knowledge of a fact, which is a function of semantic memory.

Implicit memory (non-declarative memory) is a component of long-term memory that refers to the storage of unconsciously acquired and accessible information (Stangor & Walinga, 2014). It includes information acquired by priming, skill learning, classical conditioning, and other performance-based information (Schacter & Tulving, 1994).

Priming refers to an enhanced ability to detect or recognize similar sensory stimuli recently exposed (Squire et al., 1993). For instance, in a memory task that requires remembering words, people tend to recall the words they were exposed to earlier better than the new words.

1.1.2 Schema Theory

According to schema theory, schemas are mental structures that represent an individual's prior knowledge, experiences, and expectations about a particular topic or event. These schemas help individuals process new information efficiently by quickly categorizing and interpreting based on their existing knowledge (Bartlett, 1932; Brainerd et al., 2005).

Schema theory includes various principles, but we will focus on four fundamental principles: selection, abstraction, interpretation, and integration.

The selection principle proposes that only some parts of exposed events, words, or people can be remembered. In other words, memory encoding functions as a coffee filter in which some information passes (encoded) while others stay on the filter and are discarded. Encoding information into memory depends on consistency with existing memory schemas. If the information aligns with active schemas, it will likely be processed and stored in memory

(Brainerd et al., 2005). The selection process assumes that the most consistent information of the relevant schema has privileged attention. Therefore, this information tends to be retained in memory (Brown & Smiley, 1977).

The abstraction principle proposes that the amount of detail encoded depends on the typicality of the relevant activated schema. More specifically, the level of detail of the information selected for encoding is low if the information is highly consistent with the activated schema. Therefore, when information is encoded, it tends to undergo an abstraction process if it has more prototypical features, discarding distinctive features from the information's memory representation (Brainerd et al., 2005).

Opposite to abstraction, interpretation is the principle that enriches information's details in-memory representation. When a schema is activated, information consistent with the schema and information not experienced but typical for the schema are generated and encoded into memory. Interpretation leads individuals to go beyond what they experienced (Brainerd et al., 2005).

The integration principle refers to reinforcing memory representations regarding an activated schema after storing a memory. When a memory schema is activated, the representations of the experience move to the final assimilation process that merges them into a permanent long-term memory structure. As a result, memories of experienced items retained from situations where memory schemas were activated are not isolated but instead form a unified, consolidated representation (Brainerd et al., 2005).

1.1.3 Dual-Processing Model

The dual-processing model suggests that memory processing occurs through two distinct processes: automatic and controlled. The automatic process is rapid, effortless, and unconscious, while the controlled process is slower, more effortful, and conscious (Gawronski & Creighton, 2013). Within this model, memory retrieval can occur in recognition and recall (Brainerd et al., 2005). Recognition involves automatically identifying information stored as representations of stimuli in the memory system (Tulving & Pearlstone, 1966; Tulving, 1983). When a stimulus encounters a stored representation, it is compared to pre-existing representations for a match, and if aligned, it is identified as previously encountered (Mandler, 1980).

Mandler (1980) proposed two types of recognition: familiarity-based and recollection-based. Familiarity-based recognition involves remembering the stimulus without recalling specific details, whereas recollection-based recognition involves recalling particular information about the stimulus, including where and when it was encountered (Yonelinas, 2002). In contrast, recall requires more effort than recognition and involves retrieving various cues from memory, such as the context of learning or the individual's mood or mental state (Tulving & Craik, 2000; Brainerd et al., 2005).

1.1.4 Fuzzy-trace Theory

Fuzzy-trace theory is a cognitive psychology framework that builds on principles from the dual-processing approach. According to this theory, memory is not a unified system but comprises two distinct processes: verbatim and gist traces. Verbatim traces refer to the exact details of an event or information, while gist traces capture the essential meaning or overall summary (Brainerd & Reyna, 2005). The theory posits that people rely more on gist traces than verbatim traces when recalling information. Gist traces are considered more stable and durable, whereas verbatim traces are more prone to decay and interference. This reliance on gist traces, which are less precise but more robust, plays a crucial role in how memories are formed and recalled.

Verbatim traces encode precise, detailed information about events, including specifics like dates, times, and exact sequences of actions. These traces are highly susceptible to forgetting and interference from other memories. Conversely, gist traces encapsulate the general sense, meaning, or essence of an experience. These traces are more resilient and less vulnerable to decay over time. According to Brainerd and Reyna (2002), this dual-trace system allows for efficient memory use, balancing the need for specific detail in some situations with the broader, more generalized recall in others.

Fuzzy-trace theory suggests that individuals automatically generate verbatim and gist traces during encoding. However, the retrieval process typically favors gist traces, which are easier to access and more distortion-resistant. This preference for gist over verbatim information is due to the brain's efficiency in processing and storing meaningful rather than precise details. Reyna and Brainerd (2011) argue that this reliance on gist traces aids in making inferences and understanding complex information, which is crucial for decision-making and problem-solving.

1.1.5 Self-Memory System Model

The self-memory system is a theoretical framework of memory proposed by Conway and Pleydell-Pearce (2000). According to this model, individuals prioritize memories relevant to their current self-concepts, influencing the encoding, storage, and retrieval of autobiographical information. In essence, autobiographical memories are organized and retrieved based on their relevance to the individuals' self-concepts. The self-concept comprises three levels of organization: the individual self (personal traits), the relational self (social roles and relationships), and the collective self (group memberships and cultural identities) (Conway et al., 2004).

This model suggests that the self-concept acts as a filter through which experiences are processed, with significant events more likely to be remembered if they align with an individual's self-view. For instance, memories that reinforce one's sense of identity or social roles are more likely to be retained and easily recalled. This selective process ensures that autobiographical memory is adaptive and supports the individual's ongoing sense of self.

Further, the self-memory system posits that the interaction between autobiographical memory and self-concept is dynamic. The self-concept influences which memories are encoded and how they are organized, while memories, in turn, contribute to the construction and maintenance of the self-concept. This bidirectional relationship implies that changes in self-concept can alter the accessibility and nature of autobiographical memories. For example, McAdams (2001) highlights how life narratives and personal identity evolve, reflecting changes in how individuals view themselves and their societal roles.

1.1.6 The Source-Monitoring Framework

The source-monitoring framework explains false memories as failures in monitoring, a cognitive mechanism responsible for distinguishing the origins of our memories (Johnson et al., 1993). Reality monitoring evaluates whether memories are derived from external or internal sources (actual experiences) (thoughts and imaginations). When this mechanism fails, it can create memories that do not accurately reflect reality. This theory suggests that memory retrieval involves evaluating the origin of the memory and determining whether it was experienced, imagined, or indicated by external inputs.

The framework posits that accurate reality monitoring relies on various cues, such as contextual details and sensory information, to differentiate between actual and imagined events. For example, memories of actual events typically contain more sensory and contextual details compared to imagined events, which are often less vivid and more generic. Effective reality monitoring uses these cues to verify the authenticity of memories.

However, the source-monitoring framework also recognizes that this process is not foolproof. Several factors, including the similarity between actual and imagined events and the influence of suggestive information, can disrupt reality monitoring. When individuals encounter similar actual and imagined events, they may confuse the two, leading to misattribution of the memory's source. Additionally, suggestive information, such as leading questions or misinformation, can further complicate reality monitoring by altering the perceived origin of the memory (Brainerd et al., 2005).

1.2 False Memories

Although memory has critical roles in information processing, memory is imperfect. Instead, it often leads to malfunctions that impact our memories' accuracy, such as false memories (Brainerd et al., 2005).

False memories refer to the inaccurate retrieval of events (Rodriguez & Strange, 2015), statements (Murphy et al., 2019), or words (Roediger & McDermott, 1995). However, they are considered accurate by the individuals who hold them (Fazio & Marsh, 2010). False memories can manifest in multiple forms, such as remembering details that have never occurred, selectively recalling events, or blending different information elements into a single event (Loftus, 2003; Brainerd et al., 2005).

Various factors, such as the individual's prior knowledge, expectations, beliefs, and misinformation, influence the formation of false memories. In this section, we will examine the type of false memories, the factors that lead to false memory formations, and the explanation of false memories from the view of different memory models and theories. However, since false memories are a kind of memory malfunction, we will begin the examination with the types of memory malfunctions.

1.2.1 Type of Memory Malfunctions

Memory malfunctions have been examined in various ways in the literature. For example, Schacter (2001) divided them into seven categories or “sins of memory.”. The seven “sins” or categories follow *transience, absent-mindedness, blocking, persistence, misattribution, suggestibility, and bias* (Schacter, 2001, p. 4). Although false memories mainly focus on the sins of *misattribution, suggestibility, and bias*, the rest of the “sins” are essential to understanding memory malfunctions, especially the formation of false memories.

Transience refers to the simple forgetting of information over time. The information we obtain from our experiences becomes less accurate or decays over time. It is one of the most common memory errors we experience daily. For instance, we can easily forget what we ate at lunch yesterday. On the other hand, transience functions as a deselecting process for memory. It enhances the remaining information (Schacter, 2001; Schacter et al., 2003).

Absent-mindedness refers to forgetting information due to a lack of attention during encoding. For example, someone may misplace the keys or forget someone's name due to a lack of attention when introducing themselves (Schacter, 2001).

Blocking refers to the difficulty or temporary inability to retrieve information stored in memory (Schacter, 2001). A typical example of blocking the tip-of-the-tongue (TOT) phenomenon is an experience in which an individual feels like they know a word or name but cannot retrieve it from memory. Furthermore, Schwartz (1999) showed that during a TOT state, the left inferior frontal gyrus and the anterior cingulate cortex were significantly more active than during a prosperous retrieval state.

Persistence is the opposite of forgetting. It refers to the unwanted retention of memories we would rather forget. Those memories might be extreme events drastically influencing an individual's life, such as traumatic events or negative experiences (Schacter, 2001). For instance, individuals who are diagnosed with post-traumatic stress disorder (PTSD) experience intrusive thoughts and memories of a traumatic event. Unrelated stimuli can trigger these thoughts and memories, which can cause significant distress and impairment in daily functioning (American Psychiatric Association, 2013).

Misattribution is a false memory that occurs when a memory is attributed to the wrong source. More specifically, although individuals might remember a particular piece of information, they incorrectly attribute it to a different source than its source. The source confusion might occur as the events and times are incorrectly bound. Several factors, such as suggestion, familiarity, and the passage of time, can influence incorrect attribution. For instance, we might question our experiences in a moment of uncertainty while leaving their house, such as whether we closed and locked the door. Later, while driving, we might panic as we cannot recall whether we closed or imagined the door (Schacter, 2001).

Suggestibility, or “rich false memories” (Loftus, 2003), refers to a memory malfunction generated by external factors such as leading questions or misinformation. It often leads to rich false memories, such as believing and recalling events that have not occurred. The “lost-in-the-mall” paradigm (Loftus & Pickrell, 1995) is one of the most famous examples of the influence of suggestibility on false memories in false memories literature. In this paradigm, participants are given a booklet containing an event about getting lost in a shopping mall as a child. Then, they were asked to imagine themselves as part of the story and remember the event as much as possible. The results of the study showed that a significant number of participants reported a false memory of getting lost in a shopping mall.

Bias refers to a memory malfunction that appears as misremembering information because of the influences of preexisting beliefs, attitudes, or knowledge. It plays a significant role in the formation of false memories. Bias can be observed in various forms, such as consistency, change, hindsight, egocentric, and stereotypical biases (Schacter, 2001).

Consistency bias leads individuals to reconstruct memories consistent with preexisting beliefs, attitudes, or knowledge. The impacts of consistency bias can be observed in various ways, such as selectively recalling the details of an event consistent with preexisting beliefs (Schacter, 2001).

Change bias refers to overestimating the change in individuals or others over time. It often results in the illusion of success in individuals (Schacter, 2001). Although it is not a well-studied area in memory literature, studies have shown supporting findings.

Hindsight is a bias in which individuals overestimate their ability to predict an event after it has occurred (Schacter, 2001). In other words, it is the “knew it all along” bias. For instance, individuals might suggest that they know the score before the match even starts or the outcomes of political events (Fischhoff, 1975).

Egocentric bias refers to remembering information in a self-enhancing or self-protective way (Schacter, 2001). It influences the individuals’ memory, such as misremembering themselves as performing better than they did or rejecting adverse outcomes.

Stereotypical bias shows similar features compared to consistency bias. However, stereotypical bias is a consistency bias that leads individuals to remember information in a way consistent with stereotypes or social norms (Schacter, 2001). Research has shown that stereotypical bias influences individuals’ recall of a particular group’s behavior or attitudes. For example, Marsh et al. (2003) found that participants who have negative attitudes toward African Americans were more likely to recall negative information about them.

1.2.2 Schema Theory

The principles of schema theory—selection, abstraction, interpretation, and integration—play critical roles in forming false memories. Schemas are mental structures that help individuals organize and interpret information. They enable people to quickly process complex information by relating new data to existing cognitive frameworks (Bartlett, 1932; Alba & Hasher, 1983).

The selection principle proposes that not all information received is encoded into memory. Instead, encoding focuses on details consistent with existing schemas. This selective process ensures that only the most relevant information is retained, filtering out less pertinent details (Bransford & Johnson, 1972). This selective encoding is crucial for managing cognitive load but can lead to the exclusion of essential but schema-inconsistent details.

The abstraction principle suggests that memory encoding relies heavily on the typicality of information relative to activated schemas. Highly consistent details with these schemas undergo abstraction, where less typical features are discarded, leading to generalized and sometimes inaccurate memory representations (Brainerd et al., 2005). For example, when recalling an office scene, an individual might remember typical items like desks and computers but forget

atypical items like a plant or a painting, illustrating how abstraction simplifies but can distort memory.

The interpretation principle extends the concept by proposing that when schemas are activated, not only is existing information encoded but additional, non-experienced details may also be generated based on the individual's expectations and the schema's influence. This process can enrich the memory with fabricated elements that align with the individual's expectations but were never experienced, leading to false memories (Brainerd et al., 2005). For instance, participants might falsely remember seeing books in a professor's office even if none were present due to the strong association between books and the office schema (Brewer & Treyens, 1981).

Furthermore, the integration principle highlights that once memories are stored, they are not isolated but integrated within a broader cognitive schema. This process merges actual experiences with fabricated details into a cohesive memory representation, critical for understanding how false memories are formed. This unified approach to memory storage supports the one-process hypothesis, which posits that both accurate and inaccurate memories stem from the same foundational cognitive processes (Brainerd et al., 2005). This hypothesis explains why false memories can be recalled confidently, as they are stored indistinguishably from true memories within the schema framework.

Schema theory comprehensively explains how memories are formed, retained, and sometimes distorted. These distortions, manifesting as false memories, are a natural byproduct of the brain's tendency to integrate new information into pre-existing schemas. That highlights human memory's reconstructive and fallible nature. Loftus and Palmer (1974) demonstrated that post-event information could alter memories, such as changing the verb in a question about a car accident (e.g., "smashed" vs. "hit") influenced participants' recall of the event's severity.

Alba and Hasher's (1983) research proved that schema-driven memory errors are common. They found that when participants were presented with a list of items to remember, they often recalled schema-consistent but non-presented items, demonstrating the powerful influence of schemas on memory reconstruction.

Brewer and Nakamura (1984) expanded on the idea that schemas guide encoding and retrieval processes. They showed that when individuals are asked to recall details from a previously experienced event, their memories are often reconstructed based on their existing schemas, which can lead to the inclusion of incorrect details.

Furthermore, Neuschatz et al. (2002) found that individuals are likelier to develop false memories of events that fit well with their schemas. Participants who read stories containing both schema-consistent and schema-inconsistent information were more likely to recall the schema-consistent information accurately but often falsely remembered additional schema-consistent details that were not present.

From a linguistic perspective, presenting information can also influence memory reconstruction. Pragmatics, the study of how context influences language interpretation, shows that people often fill in gaps in their understanding based on their schemas (Grice, 1975). This pragmatic inference means that people are likely to "remember" implied details but not explicitly stated, aligning with their existing schemas (Wilson & Sperber, 2012).

Brewer (1977) and Nieuwland and Van Berkum (2006) found that pragmatic inferences drawn from conversations often become integrated into memory, affecting recall accuracy. This integration demonstrates that linguistic factors, such as how information is conveyed and interpreted, significantly impact how memories are reconstructed.

Schema theory also provides a framework for understanding how just-world beliefs can impact memory. Just-world beliefs act as schemas that guide the encoding and retrieval of information. When individuals encounter information that aligns with their just-world beliefs, it is more likely to be encoded and remembered accurately. Conversely, information contradicting these beliefs may be distorted or forgotten to maintain a consistent worldview.

Hafer and Bègue (2005) found that individuals with strong just world beliefs are more likely to distort memories to fit their beliefs, particularly when faced with evidence of injustice. This finding aligns with schema theory's abstraction and integration principles, where inconsistent details are discarded or altered to fit the existing schema.

Furthermore, Dawtry et al. (2019) showed that individuals with strong just world beliefs are prone to blaming victims for their misfortune, which involves reconstructing memories to align with the belief that the world is fair. This cognitive bias supports the integration principle, as new information is modified to fit the pre-existing just world schema.

Callan et al. (2009) demonstrated that people with strong just world beliefs selectively remembered personal deeds that aligned with their own good or bad fortune. This selective memory process supports the selection principle of schema theory, where only schema-consistent details are encoded and retained.

By understanding the role of schemas in memory and the influence of just world beliefs, we can better appreciate how cognitive biases shape our recollections and perceptions. This integration of schema theory and the just world hypothesis highlights the complex interplay between cognitive structures and belief systems in shaping memory.

1.2.3 The Source-Monitoring Framework

The Source-Monitoring Framework (SMF) provides a framework for understanding how false memories can arise from source-monitoring errors. When individuals retrieve a memory, they must assess the source attributes and judge its origin. If this process is flawed, false memories can be created. For example, an individual might recall a vivid dream and, due to a lack of explicit source attributes, believe it to be an actual event. Similarly, misinformation introduced after an event can be misattributed to the original event, leading to false memories (Loftus & Hoffman, 1989).

The SMF can also be applied to understand how world beliefs influence memory recall. Individuals with strong just world beliefs may be more prone to source-monitoring errors when the information contradicts their belief in a fair world. For instance, they might misattribute negative information about a perceived "good" person to an unreliable source, thereby maintaining their belief in a just world.

Linguistic cues play a significant role in source monitoring. The way information is presented can influence how it is remembered and attributed. As studied by Grice (1975) and Wilson and Sperber (2012), Pragmatic inferences show that individuals often rely on contextual clues to

infer the source of information. It means that language and the context in which it is used can lead to source-monitoring errors if the cues are misleading or ambiguous.

Brewer (1977) and Nieuwland and Van Berkum (2006) found that the pragmatic inferences drawn from conversations often become integrated into memory, affecting recall accuracy. It demonstrates that linguistic factors, such as how information is conveyed and interpreted, significantly impact how memories are reconstructed and attributed to sources.

The source-monitoring framework provides a comprehensive understanding of how memories are attributed to their sources and how errors can lead to false memories. By integrating cognitive psychology and linguistics findings, we can better appreciate the complexity of memory reconstruction and the potential for bias in source monitoring. Understanding the role of just world beliefs in this framework further highlights the interplay between cognitive processes and belief systems in shaping memory.

1.2.4 Dual-Process Model

According to the dual-processing model, the interaction between Systems 1 and 2 can lead to memory distortions, primarily when System 1 generates an initial, automatic response that System 2 fails to verify or correct adequately. For instance, in eyewitness testimony, an initial impression formed by System 1 might include incorrect details due to suggestion or misinformation. If System 2 does not intervene to correct these errors, the false details can become part of the individual's memory (Loftus, 2005).

Roediger and McDermott (1995) demonstrated the role of associative processes in false memory formation. In their study, participants were presented with lists of semantically related words and later tested for recall and recognition. Many participants falsely recalled and recognized non-presented words strongly associated with the presented words, illustrating how System 1's associative processes can lead to false memories.

In another study, Gallo et al. (2001) explored how warnings about false memories could influence recall accuracy. Even when participants were warned about the potential for false memories, many still experienced them, indicating that System 1's automatic processes are robust and not easily overridden by System 2's controlled processes.

The dual-process model also provides insights into how just world beliefs can influence memory. Individuals with strong just world beliefs may rely more heavily on System 1's automatic processing to maintain their belief in a fair world. System 1 might automatically generate justifications or distortions to align with the just world schema when confronted with information that contradicts this belief. In contrast, System 2 might be less effective in correcting these distortions due to motivational biases.

Hafer and Bègue (2005) explored the interplay between just world beliefs and cognitive processing. They found that individuals with strong just world beliefs are more likely to engage in biased processing that favors their belief system, which can lead to memory distortions. This supports the idea that System 1's automatic processes, influenced by just world beliefs, can dominate memory processing, with System 2 failing to correct these biases effectively.

Another study by Lerner and Simmons (1966) showed that when participants observed others suffering, those with strong just world beliefs were more likely to derogate the victim to maintain their belief in a fair world. This suggests that just world beliefs can influence the initial encoding of information and its later recall, as System 1 automatically generates justifications that System 2 may not adequately challenge.

Linguistic cues also play a crucial role in the dual-process model of memory. How information is presented can trigger System 1's automatic processes or System 2's controlled processing. As studied by Grice (1975) and Wilson and Sperber (2012), pragmatic inferences can influence how information is encoded and recalled. For example, implied meanings and context can lead to automatic inferences that shape memory recall, often without the individual being consciously aware of these influences.

Consequently, the dual-process model provides a cognitive psychology and linguistics framework for understanding how memories are formed, retained, and sometimes distorted through the interaction of automatic and controlled cognitive processes.

1.2.5 Fuzzy-trace Theory

According to the Fuzzy-trace Theory (FTT), false memories often rely on gist rather than verbatim memory. When individuals recall an event, they may rely on the gist representation to

reconstruct the memory, which can lead to distortions if the gist is incorrect or incomplete. This process can include details that fit the general theme but were not part of the original event (Brainerd & Reyna, 2005).

Seamon et al. (2002) demonstrated that false memories can be created by suggesting semantically related but non-presented words. Participants given lists of semantically related words often falsely recalled words that were not presented but fit the gist of the list. This supports the idea that gist representations can lead to false memories.

Another study by Tun et al. (1998) investigated memory in older adults and found that they are more likely to rely on gist memory than younger adults, leading to a higher incidence of false memories. This finding supports the notion that reliance on gist memory can result in memory distortions, especially when verbatim memory is weak.

FTT also provides a framework for understanding how just world beliefs can influence memory recall. Individuals with strong just world beliefs may rely more heavily on gist representations that align with their belief in a fair world. When recalling events, these individuals might reconstruct memories based on the general theme of justice or fairness, leading to distortions that support their just-world beliefs.

A study by Tobin and Raymundo (2009) found that individuals with strong just-world beliefs were more likely to recall information in a way that supported their belief system. This suggests that gist representations influenced by just world beliefs can dominate memory recall, leading to biased reconstructions of events.

Another study by Callan et al. (2007) showed that people with strong just world beliefs are likelier to remember information that supports their beliefs and forget information that contradicts them. This finding aligns with FTT, demonstrating that gist memory, influenced by just-world beliefs, can lead to selective memory retention and recall.

Linguistic cues also play a significant role in fuzzy-trace theory. How information is presented can influence whether verbatim or gist representations are formed. As studied by Grice (1975) and Wilson and Sperber (2012), pragmatic inferences can shape how information is encoded and recalled. For example, implied meanings and context can lead to gist representations that shape memory recall, often without the individual being consciously aware of these influences.

Research by Murphy and Shapiro (1994) demonstrated that how sentences are framed can influence whether verbatim or gist memory is used. They found that pragmatic inferences drawn from conversations are integrated into memory, demonstrating that linguistic factors significantly impact how memories are formed and recalled.

Another study by Christianson and Loftus (1991) showed that misleading post-event information could alter memory by affecting gist rather than verbatim representations. This indicates that the linguistic presentation of information can significantly impact memory accuracy.

Fuzzy-trace theory provides a robust framework for understanding how memories are formed, retained, and sometimes distorted through the interaction of verbatim and gist representations. By integrating cognitive psychology and linguistics findings, we gain a deeper appreciation of how both types of memory processing contribute to memory reconstruction and the potential for bias. Understanding the role of just world beliefs within this framework further underscores the interplay between cognitive processes and belief systems in shaping memory.

1.2.6 Self-Memory System Model

The Self-Memory System Model (SMS) model also explains how false memories can form. Since current goals and self-concept influence the retrieval of autobiographical memories, there is a potential for distortion. Memories may be altered or fabricated to fit the current narrative of the self-better. This can lead to the creation of false memories that are consistent with one's self-image and present concerns (Conway, 2005).

A study by Mazzoni and Memon (2003) demonstrated how the SMS model could lead to false memories. Participants were led to believe they had experienced a false childhood event (e.g., getting lost in a shopping mall). Those with strong motivations to accept the event as part of their self-narrative were likelier to develop false memories of the event, supporting the idea that the working self influences memory construction.

Another study by Wilson and Ross (2003) found that people recall past events to enhance their current self-view. Participants often remembered past selves as less competent or successful

than present ones, illustrating how the working self shapes memory to support a positive self-concept.

The SMS model also provides insights into how just world beliefs influence memory recall. Individuals with strong just world beliefs may selectively retrieve and construct memories that align with their belief in a fair world. This selective retrieval can reinforce their just-world beliefs by filtering out memories contradicting this worldview and emphasizing those supporting it.

A study by Dalbert (2001) found that individuals with strong just world beliefs often recalled events in a way that supported their belief system. For example, they were more likely to remember instances where good things happened to good people and bad things happened to bad people, aligning their memories with their just-world beliefs.

Similarly, a study by Anderson and Conway (1993) showed that autobiographical memories are often reconstructed to fit current self-concepts and belief systems. Participants with strong just world beliefs were found to reconstruct their memories to align with these beliefs, further supporting the SMS model's assertion that the working self influences memory retrieval.

Linguistic cues also play a significant role in the self-memory system model. How information is framed can influence how it is integrated into the autobiographical knowledge base and retrieved by the working self. As Grice (1975) and Wilson and Sperber (2012) studied, pragmatic inferences can shape how autobiographical memories are encoded and recalled. For example, implied meanings and contextual cues can lead to the selective retrieval of memories that support the current self-concept.

A study by Pasupathi and Wainryb (2010) demonstrated that people's narratives about their lives are influenced by their current self-concept and the social context in which they are recalled. This suggests that linguistic factors significantly impact how autobiographical memories are formed and recalled, supporting the SMS model's emphasis on the dynamic interaction between the self and memory.

Another study by Edwards and Middleton (1986) found that the language used in autobiographical narratives can shape memory recall. Participants' memories were influenced

by how they discussed past events, illustrating how linguistic framing can impact the integration of experiences into the autobiographical knowledge base.

The self-memory system model provides a robust framework for understanding how autobiographical memories are formed, retained, and sometimes distorted through the interaction of the autobiographical knowledge base and the working self. By integrating cognitive psychology and linguistics findings, we gain a deeper appreciation of how these processes contribute to memory reconstruction and the potential for bias.

1.2.7 Neurobiological Perspectives of False Memories

Recent advancements in neurobiological research have shed light on the intricate mechanisms underlying false memories, revealing a fascinating interplay between various brain regions and cognitive processes. These insights not only enhance our understanding of memory distortions but also pave the way for developing strategies to mitigate the impact of false memories.

The hippocampus, a critical region for forming and retrieving episodic memories, has been a focal point in studying false memories. Through innovative techniques like optogenetics, researchers have demonstrated that activating specific groups of hippocampal cells can induce the recall of false memories. In groundbreaking experiments, scientists could implant false fear memories in mice by stimulating hippocampal neurons associated with a particular context. This highlights the hippocampus's role in both the formation of true memories and the creation of false ones, illustrating how memory traces can be manipulated (Tonegawa et al., 2015).

Furthermore, the prefrontal cortex is crucial in constructing and retrieving false memories. This region is responsible for higher-order executive functions such as planning, decision-making, and integrating various pieces of information. Increased activity in the prefrontal cortex during false memory retrieval suggests that this area is heavily engaged in the cognitive effort required to reconcile conflicting information and construct plausible but inaccurate recollections (Schacter & Slotnick, 2004). This involvement underscores the brain's remarkable capacity for creating coherent narratives, even when the details are fabricated.

Synaptic plasticity, the process by which synaptic connections are strengthened or weakened, is fundamental to memory formation. Long-term potentiation (LTP), a mechanism that strengthens

synaptic connections through repeated stimulation, is crucial for embedding both true and false memories. This process shows how repeated exposure to certain stimuli can lead to the integration of false information into memory networks, demonstrating the brain's ability to adapt and reorganize itself based on experiences (Neves et al., 2008). The plastic nature of synapses highlights the dynamic aspect of memory, where the exact mechanisms that enable learning and memory can also facilitate the formation of false memories.

The amygdala, known for processing emotions, also plays a significant role in false memories. Research shows heightened activity in the amygdala during the recall of emotionally charged false memories, indicating that emotional arousal can enhance the formation and retrieval of these memories. This suggests that the emotional context in which memories are formed can significantly influence their accuracy and vividness (Schacter et al., 1996). The interplay between the hippocampus and amygdala underscores the importance of emotional factors in memory distortion.

Age-related changes in the brain can also impact susceptibility to false memories. Older adults are more prone to false memories due to alterations in frontal and temporal lobes, which affect their ability to encode and retrieve detailed information accurately. These changes highlight the impact of aging on memory processes and the increased cognitive load required to maintain memory accuracy in older individuals (Dennis et al., 2007).

Moreover, memory consolidation during sleep has also been linked to false memories. Research indicates that sleep can consolidate both true and false memories, suggesting that the mechanisms involved in memory consolidation do not discriminate between actual and imagined events. This finding emphasizes the need to explore how sleep influences memory accuracy and the potential for false memories to become entrenched (Diekelmann et al., 2009).

In summary, the neurobiological underpinnings of false memories involve a complex network of brain regions and processes. From the hippocampus's role in encoding and retrieval to the prefrontal cortex's involvement in cognitive control and the amygdala's influence through emotional arousal, false memories are the product of an intricate web of neural activities.

1.3 Motivated Memory

The standard view of the motivations influencing cognition is usually based on striving for pleasure and avoiding pain (Higgins, 2012). Although this view of motivation covers a wide range of motivational influences on cognition, motivational factors can vary in specific components of cognition, such as memory (Eitam et al., 2013). Numerous studies have shown that sensory and short-term memory encoding influences motivations and goals. For instance, individuals tend to perceive image interpretations in a way consistent with their benefits (Balceris & Dunning, 2006). On the other hand, the influence of motivation on selectivity could be observed in all cognitive processes, such as decision-making. For example, individuals are motivated to identify fake news that supports their opposite ideological beliefs as fake more than fake news that supports their ideological beliefs (Murphy et al., 2019).

This motivational influence on memory affects not only what is remembered but also what is forgotten or misremembered. This leads us to motivated forgetting and misremembering, where individuals actively or passively alter their memories to maintain psychological equilibrium.

1.3.1 Motivated Forgetting and Misremembering

Motivated forgetting refers to the motivational inhibition of perceived as threatening or distressing memories. It is a coping mechanism in which individuals suppress the memories that threaten their prior thoughts about themselves and the groups they belong (Tajfel & Turner, 1986). The motivational aspect of this inhibition is examined from various perspectives in the memory literature. However, we can examine it through two categories: *neglected encoding and retrieval* (Dalton & Huang, 2014).

Neglected encoding refers to the motivation of neglecting the experiences that threaten individuals or their groups in the memory encoding process. It allows individuals to cope with self-concept-threatening experiences such as traumas, which results in reduced memory (Sedikides et al., 2004).

According to Freud (1953), individuals engage in the forgetting process to protect their self-views from threatening memories. Based on this, neglected retrieval is a forgetting that blocks

encoded information from being transferred to consciousness and, consequently, the retrieval process (Dalton & Huang, 2014).

On the other hand, forgetting is not the only way individuals resort to coping with self-threatening experiences. Instead, individuals are sometimes motivated to modify their memories to align with one's beliefs or goals. This leads to biased retrieval, also known as motivated memory recalling. For instance, low self-esteem individuals tend to remember their negative memories as more optimistic than they were to maintain a positive self-image, which has been called the "fading affect bias" (Markus & Wurf, 1987; Pyszczynski & Greenberg, 1987; Skowronski et al., 1991; Walker & Skowronski, 2009).

Fading affect bias is a kind of motivated misremembering where people tend to remember negative emotions less accurately over time than positive ones. According to Walker et al. (2003), it is a psychological mechanism that allows individuals to cope with negative experiences.

Additionally, various factors can impact motivated forgetting and misremembering. For instance, individual differences significantly affect the tendency to engage in motivated forgetting and misremembering. Individuals with high levels of neuroticism are more likely to engage in memory suppression. In contrast, individuals with high levels of self-esteem are more likely to engage in memory distortion (Mather & Nesmith, 2008).

Furthermore, situational factors can also influence the tendency to engage in motivated forgetting and misremembering. Individuals are more likely to suppress memory when exposed to highly emotional or traumatic information (Brewin, 2007). Additionally, the type of information and the context in which it is presented can influence the tendency to engage in motivated forgetting and misremembering.

Cultural norms can also influence the tendency to engage in motivated forgetting and misremembering. For example, collectivist cultures, where interpersonal harmony is valued over individual expression, tend to engage in motivated forgetting and misremembering to maintain social harmony (Mather & Nesmith, 2008).

1.3.2 Cognitive Dissonance

Cognitive dissonance is a psychological phenomenon first proposed by Festinger in 1957. As a novel approach, he considered cognition a piece of knowledge with a psychological representation. An individual might have unlimited cognition interacting with or completing each other. Individuals strive for inner consistency (consonance) among those cognitions (Festinger, 1968, p. 1). However, the cognitions sometimes contradict each other. More specifically, when one cognition follows from the opposite of another, the pair of cognitions is considered inconsistent (dissonant). For instance, an individual with negative stereotypes toward an ethnic group tends to develop negative thoughts and attitudes toward a neighborhood where the ethnic group usually lives, such as the crime rate would be high. However, inconsistent with the individual's negative stereotypes, the neighborhood where the ethnic group lives might show low crime rates. In such a case, two cognitions are inconsistent or dissonant because the low crime rates of the ethnic group's neighborhood contradict her stereotypes. This dissonance between two cognitions leads individuals to experience uncomfortable tension, also known as psychological discomfort. Furthermore, the discrepancy between the two cognitions defines the magnitude of psychological discomfort. Individuals who experience more significant psychological discomfort are more motivated to reduce it.

The psychological discomfort motivates individuals to attempt to rationalize the inconsistency between two cognitions. Regarding our example, the attempts could appear as follows: the ethnic group will eventually show high crime rates, or this neighborhood is an exception. However, individuals' rationalization attempts are not always successful. Instead, they might fail, and the dissonance continues (Festinger, 1968). The dissonance between two cognitions where the rationalization attempts failed motivates individuals to change the psychological representations of those cognitions. The changes could occur in two ways:

- 1) Individuals change their prior cognition about their attitudes, thoughts, or beliefs consistent with the other cognition.
- 2) Individuals change their cognition about the psychological representation, which contradicts their prior cognition (Festinger, 1968, p. 6).

Considering our example, the individual changes her negative stereotypes about the ethnic group with positive traits or distorts or modifies the information about the crime rate in the neighborhood where the ethnic group lives.

In addition, brain imaging studies showed that cognitive dissonance in individuals is associated with the dorsal anterior cingulate cortex (ACC), which is often related to cognitive conflict; the left anterior insula (left aINS), which is related to aversive emotional arousal; and the posterior cingulate cortex (PCC), which is connected to self-processing (Kitayama et al., 2013; Colosio et al., 2017).

1.3.3 Belief in a Just World

Our environment contains many unpredictable variables, such as weather and earthquakes. In addition to natural uncertainties, it includes social and individual unpredictable variables such as inflation, wars, success, and rewards. On the other hand, individuals need some assumptions about themselves, the world around them, and the future to achieve long-term goals, understand the world, and build their self-concepts (Lerner, 1980). For instance, students assume that their scores and efforts on an exam correlate to their desire to achieve high scores.

The belief in a just world theory suggests that individuals are motivated to believe in a fair world where people get what they deserve. More specifically, individuals deserve their fortune or misfortune if they have met the conditions to achieve this fortune or misfortune. This belief system allows individuals to predict the future, build their self-concepts, and achieve long-term goals (Lerner, 1980; Dalbert, 1998). For example, an individual who works harder to get promoted to a higher position in the company assumes that their efforts and dedication have been recognized and rewarded. More specifically, the individual believes that they will get what they deserve. This belief allows them to maintain a positive outlook on their future in the company and build a strong self-concept as a competent and valuable employee.

Every new information entered into cognition is initially coded and compared with the individual's pre-existing knowledge, also known as cognitive templates. Cognitive templates also refer to the base of an individual's belief system. Like other beliefs, an individual's belief in a just world must be based on a cognitive template fed by some information (Lerner, 1980).

The information that feeds the just world belief can be examined into two categories: personal experiences and social influence.

Personal experiences can influence an individual's belief in a just world, mainly when justice is ambiguous or unclear (Lerner, 1980). For instance, an individual who has always lived a comfortable and privileged life likely experienced a world they perceive as just and fair because they have never witnessed any major injustices or suffered significant hardships.

Morality tales are stories that are passed down through generations. They often center around themes of justice and fairness and teach a moral lesson or convey a particular message about right and wrong (Lerner, 1980). Social norms are the unwritten rules representing suitable behaviors or thoughts for a specific culture. Individuals learn and emphasize those rules from social interactions or their parents. The socialization process in which individuals learn social norms and values can shape an individual's belief in a just world. Studies have shown that socialization processes emphasizing fairness, such as religious or moral teachings, can lead to a stronger belief in a just world (Dalbert & Sallay, 2004).

Social status refers to an individual's position in a social hierarchy, such as wealth, occupation, or education level. Individuals with higher status tend to have stronger beliefs in a just world, while those with lower status tend to have weaker beliefs. Those with higher status tend to have more control over their lives; thus, they are less likely to experience injustice. They also have more resources to cope with injustice. In addition, social status is also related to social norms the individual supports. Higher-status individuals often support hard work, individualism, and meritocracy. In contrast, lower-status individuals may endorse values such as collectivism and social justice, emphasizing the need for reducing inequality and promoting fairness in society (Dalbert & Sallay, 2004).

Religion can be essential in shaping an individual's belief in a just world. In many religions, individuals are responsible for their actions, and injustice will be served in the afterlife. As a result, religious individuals often believe in a just world more than non-religious people. Additionally, religion may provide individuals with a framework for interpreting and coping with injustice, which can help to reinforce beliefs in a just world (Dalbert & Sallay, 2004).

Although individuals need a just world where they can predict future outcomes, the reality of injustice might contradict their prior beliefs. This inconsistency creates conflict in individuals' cognition, called cognitive dissonance. Cognitive dissonance leads to psychological discomfort, which individuals are motivated to reduce (Lerner, 1980). Instead of modifying their beliefs about the world as a fair place, individuals tend to resort to some strategies to reduce psychological discomfort (Lerner, 1980). We can examine those strategies into rational and irrational strategies (Lerner, 1980).

Prevention and restitution are two of the most common rational strategies for reducing discomfort caused by injustice. They refer to all kinds of aids to prevent the situations that cause injustice (Lerner, 1980). For instance, an individual can donate money to infant orphans to justify their belief in a just world and reduce the psychological discomfort the injustice causes.

On the other hand, individuals can accept their limitations in situations that cause injustice (Lerner, 1980). For instance, an individual might donate money for infant orphans but accept that they cannot help all the orphans in the world.

The rational attempts to reduce psychological discomfort sometimes might not be enough. Under such a consequence, individuals resort to irrational attempts to reduce psychological discomfort and justify their beliefs in a fair world.

Denial-withdrawal is a primitive defense mechanism that consciously selects the information to which one is exposed. It helps individuals maintain a sense of psychological comfort without directly distorting reality or facts. For instance, individuals may avoid visiting poor areas, emergency rooms, or institutions to shield themselves from witnessing suffering or injustice. Even if they experience distressing events, they quickly withdraw physically and mentally.

Reinterpreting the event involves various strategies to reinterpret perceived injustices, thereby minimizing or eliminating distress. One common strategy is to reinterpret the outcome, framing the victim's suffering as beneficial or desirable. For example, individuals may rationalize poverty or hardship as character-building experiences that lead to personal growth or spiritual enlightenment. Religious beliefs may reinforce this interpretation by promising restitution or rewards in the afterlife for enduring suffering in the present (Lerner, 1980).

Another strategy involves reinterpreting the cause of the injustice. Individuals tend to view the victim's suffering as their actions or shortcomings. This attitude allows individuals to rationalize the suffering of others and avoid feelings of guilt or responsibility. In specific cases, individuals even feel satisfied with the belief that victims "deserve" their misfortune.

Additionally, individuals may reinterpret the victim's character to justify their suffering or disadvantaged status. Socially constructed stereotypes and prejudices contribute to the normalization of inequality and suffering for specific groups. Regarding this strategy, those members of disadvantaged communities are in disadvantaged positions because of their laziness and deserve their misfortunes and suffering (Lerner, 1980).

Various studies have supported these irrational strategies. For instance, blaming the victim of a mugging for being careless with their belongings leads to less support for social justice policies to prevent similar incidents (Dawtry et al., 2019). Furthermore, individuals recalled a smaller lottery prize when the winner was 'bad' and selectively remembered personal deeds that aligned with their good or bad fortune (Callan et al., 2009).

Recent studies have further connected just world beliefs with memory processes. Hafer and Bègue (2005) found that individuals with a strong belief in a just world are more likely to distort memories to fit their beliefs, particularly when faced with evidence of injustice. This aligns with the broader literature on motivated memory, where individuals' desires to see the world as fair lead to selective memory processes (Hafer & Bègue, 2005).

Moreover, memory recall and false memory formation can be influenced by linguistic factors. Research by Grice (1975) and Wilson and Sperber (2012) on implicature suggests that individuals often rely on pragmatic inferences to fill gaps in memory. This means that just-world beliefs could influence how people reconstruct memories, as they might be more inclined to infer details that align with their belief in a fair world.

Integrating linguistics and memory, Brewer (1977) and Nieuwland and Van Berkum (2006) found that the pragmatic inferences people draw from conversations often become integrated into their memory, affecting recall accuracy. This supports the idea that individuals with a strong belief in a just world might reconstruct memories using implicatures that support their beliefs.

In a related study, Greene and Loftus (1984) demonstrated that people are prone to reconstructing their memories in ways that align with their current beliefs and attitudes. This can be seen as a cognitive bias that just world believers might use to maintain their worldview.

Understanding the role of just world beliefs in memory recall and how these beliefs interact with linguistic processes helps to appreciate the complexity of memory formation and the potential for bias in reconstructive memory processes. This integration highlights the need for further research into how these cognitive and linguistic factors shape our recollections and beliefs.

Measurement of Belief in a Just World

In recent decades, belief in a Just World Theory has evolved with new findings and approaches. In parallel, new measurement tools have also been developed to assess various aspects of belief in a just world.

Rubin and Peplau's Just World Scale (1975) is one of the first scales developed as a measurement tool for belief in a fair world. It provides a measure of the extent to which people believe that the world is a fair place. The scale contains 20 statements (9 unjust and 11 just) and a six-point Likert scale (1 strongly disagree, and 6 strongly agree) that participants rate their agreement of statements. The Just World Scale score is calculated by subdividing the sum of the agreement degrees to the just and unjust statements.

However, previous versions of the measurement methods of belief in a just world, such as Rubin and Peplau's (1975) Just World Scale, have been criticized for their simplicity and inability to assess belief in a just world in different dimensions (Furnham & Procter, 1989). Therefore, the just world scale was revised to gain the ability to assess the extent of belief in a fair world in multiple dimensions.

One of the first examples of multidimensional belief in a just world scale is the general belief in a just world scale developed by Lipkus (1991) to address the limitations of previous scales developed for measuring belief in a just world. He suggested they were overly specific and unable to capture the nature of belief in a just world theory. Following this, Lipkusa et al. (1996) suggested that the perception of injustice differs between the situations concerning individuals and others. Furthermore, those with different behaviors and attitudes on perceiving, judging, or

denying fair and unfair situations should be assessed separately. Therefore, they suggested that the belief in a just world should be considered into two subcomponents: the belief that the world is a just place and one's actions lead to the consequences (BJW-self), and the world is a just place for others (BJW-others). For this purpose, they developed 11-item BJW-self and 11-item BJW-others scales.

Lastly, Lucas et al. (2011) suggested that besides the distinction of who experiences justice, how justice is experienced, such as procedural and distributive justice, should be measured separately. Distributive justice refers to fairness of the outcomes, allocations, or resource distribution. On the other hand, procedural justice refers to fairness in the decision-making process or the treatment. In addition, procedural and distributive justice are related to well-being and social attitudes. For instance, distributive justice is more likely related to increased positive effects, and procedural justice is more likely related to minimizing adverse effects. Therefore, they developed a new scale that covers four aspects of just world belief: distributive and procedural justice beliefs for self and distributive and procedural justice beliefs for others.

1.3.4 Accountability

Accountability refers to being responsible for an action or "answerability." However, the definition of accountability differs regarding who performs the action and who is responsible for it. More specifically, accountability shows different cognitive processes and motivations in cases where an individual is responsible for an action herself or holds others responsible for their actions (Frink & Klimoski, 2004).

The self-accountability factor leads to more careful encoding and retrieval of information, which reduces the chances of errors such as misattribution or suggestibility (Wright & Loftus, 1999). Self-accountability heightens the accuracy of memory recall by encouraging careful encoding and retrieval, thus reducing errors like misattribution and suggestibility. This phenomenon is evident in how people manage their memories and take responsibility for their accuracy, a concept referred to as "autonoesis," which denotes the self-awareness of remembering from a personal perspective (Mahr & Csibra, 2018). Conversely, social accountability involves detecting and remembering individuals who breach social contracts, as the cheater detection theory describes. This evolutionary perspective suggests that humans have developed cognitive

mechanisms to identify and remember cheaters to avoid future exploitation, thus enhancing social cooperation and survival. Such mechanisms are crucial in maintaining the integrity of social exchanges, as individuals rely on episodic memory to assess the trustworthiness of others based on their past behaviors. This dual role of accountability underscores its importance in both personal cognition and social interactions (Cosmides & Tooby, 1992).

According to the cheater detection theory, social cooperation among individuals, based on reciprocity intentions, has been an essential concept for survival in human evolution. Therefore, individuals tend to return the favors they have already received or will be received. However, these favor trades do not always work correctly. Instead, some group members might exploit this tendency by not paying the received favors back or avoiding helping the group members. That violation of the cheaters might end up with profound consequences that threaten the individual's survival. This survival threat motivates individuals to remember the others who violated the cooperation and hold them accountable for potential future violations. For instance, individuals are motivated to recognize the cheaters' faces more (Bell & Buchner, 2012), recall aggressive people less likable (Kroneisen et al., 2015), and recall unpleasant contexts more accurately (Mattarozzi et al., 2015).

False memories often arise from suggestibility and misattribution, mitigated by self-accountability. However, in social contexts, the need to remember violations of social contracts can lead to the reconstruction of memories, potentially introducing biases. This dual aspect of accountability, impacting personal cognition and social interactions, underscores its importance in memory accuracy and the potential for false memories. For instance, the belief in a just world can cause individuals to reconstruct memories to align with their need to see the world as fair (Lerner, 1980), which intersects with how accountability shapes memory.

The concept of accountability integrates seamlessly with other sections of the thesis, particularly the discussions on the belief in a just world and motivated memory recall. Just as individuals reconstruct memories to align with their belief in a just world, accountability influences memory reconstruction to maintain social cooperation and personal integrity. The interplay between cognitive and motivational factors in memory processes is evident in how accountability shapes memory accuracy and the potential for false memories. For example, when individuals believe strongly in a just world, they might be more motivated to remember or even create false

memories that support this belief, especially when their sense of fairness is challenged (Dalbert, 1998).

By understanding the role of accountability in memory processes, we can better appreciate the complexity of memory formation and recall. This highlights the need to consider cognitive and social factors carefully in research on false memories. The dual role of self-accountability and social accountability provides a comprehensive view of how personal responsibility and social interactions influence memory, thereby contributing to the broader understanding of memory accuracy and the potential for distortions in recall (Wright & Loftus, 1999; Cosmides & Tooby, 1992; Mahr & Csibra, 2018). This integration underscores the importance of accountability in both personal cognition and social interactions, suggesting that future research should continue to explore these interconnected dynamics to unravel the complexities of human memory.

1.4 Linguistics and Memory

Humans have mastered language over thousands of years as a complex communication system. While language is often taken for granted, it is a remarkable feat of human cognition that allows us to convey meaning through speech and writing (Cherry, 1966). Language study is an interdisciplinary field encompassing various aspects, such as semantics and pragmatics. In this context, linguists have identified three linguistic units essential for understanding language: word meaning, sentence meaning, and utterance (speaker's) meaning (Kroeger, 2023).

Word meaning refers to the meaning of individual words or lexical items in a language. This meaning is typically determined by the semantic features of the word, such as its category, attributes, and relationships with other words. For example, "cat" refers to a small, carnivorous mammal often kept as a pet.

Sentence meaning, on the other hand, refers to the meaning of a sentence, which is often more complex than the sum of its word meanings. It considers the sentence's syntax and how the words are combined to convey a specific message or proposition. For example, "The cat is sick" conveys information about the cat's current health condition.

Finally, *utterance (speaker's) meaning* refers to the meaning that a speaker intends to convey through a particular utterance or speech act. This meaning may be affected by various factors,

such as the context of the conversation, the speaker's tone of voice, and the speaker's intentions. For example, the utterance "Can you pass the salt?" may have different meanings depending on the context, such as whether the speaker is making a request or giving an order.

1.4.1 Semantics and Pragmatics

Semantics and pragmatics are two essential aspects of linguistics that help us understand the human language. Semantics studies how linguistic form relates to meaning and how we interpret what speakers intend to communicate. It deals with the inherent sense of words and sentences as linguistic expressions. Studying the shared system of rules allows people to communicate effectively and interpret language accurately.

Pragmatics is a branch of linguistics that deals with the study of language use in context. It examines how speakers use language to convey meaning in different social situations, considering factors such as the speaker's intentions, the listener's interpretation, and the cultural background of both parties. Pragmatics is concerned with how people use language to achieve their communicative goals and how these goals influence the meaning of what is said (Kroeger, 2023)

1.4.2 Implicature

The meaning is not always explicitly delivered in communication. It is sometimes beyond what is explicitly communicated. This aspect of communication is defined as Implicature (Sperber & Wilson, 1995). Parallel to that, the degree of explicitness determines the extent of implicature in communication.

The contribution of decoding and pragmatic inference determines the degree of explicitness. More specifically, when the speaker's meaning is Implicature, the contribution to pragmatic inference is more significant than that of decoding, which provides more cues about the speaker's intentions and context. (Grice, 1975; Wilson & Sperber, 2012).

For instance, Tom and Josh are roommates. Tom makes coffee and asks whether Josh wants to have it. He asks to Josh:

“Do you want some coffee?” Josh replies:

“Coffee would keep me awake.”

In this example, Josh’s utterance does not answer Tom’s question directly. Instead, it provides extra information about Josh and allows Tom to infer Josh’s implicatures, such as he does not want to stay awake, does not want coffee, will not drink coffee today, or wants to sleep. (Sperber& Wilson, 1995).

Although hearers must perform more cognitive processes to reconstruct the pragmatic inferences regarding the speaker’s statement, it provides cues about the speaker’s traits, context, and past and future predictions (Krauss & Fussell, 1996). Considering costs and benefits, retaining the pragmatic inferences of a speaker’s utterances is usually more profitable for individuals than the literal meaning of the utterances (Brewer & Lichtenstein, 1975). Besides, generating an inference leads to the working memory simultaneously holding multiple pieces of information in mind, enhancing the memorability of the pragmatic inferences (Nieuwland & Van Berkum, 2006). As a result of enhanced memorability, the pragmatic inferences occur in solid interaction with long-term memory, specifically episodic memory, leading individuals to recall the speaker’s utterances with their implicatures (Brewer, 1977).

The influence of time on the implicature of memory can be seen in how memories are reconstructed based on available information and previous experiences. The reconstruction of memories can be influenced by various factors, including expectations, biases, and prior knowledge (Brewer & Treyens, 1981).

For example, a person’s expectations and prior knowledge can influence the implicature of a memory. A study by Brewer and Treyens (1981) found that participants remembered more objects that were consistent with their expectations of a typical office than inconsistent objects, even if those objects were present in the room. It demonstrates how the implicature of memory can be influenced by prior knowledge and expectations.

1.4.3 Influence of Time

Memories are not stored as exact replicas of past events but rather as reconstructions based on available information and previous experiences (Schacter, 2001). The time differences between the first exposure and recall of utterances impact the retained information. More specifically,

the retained information fades or distorts over time. The Implicatures of the speaker's utterances provide more cues and less effort during retrieval. Consequently, retained utterances tend to distort and retrieve their Implicatures (Polišenská et al., 2014; Abadie & Camos, 2019).

Numerous studies have shown supportive results on the influence of time on Implicature retrieval. For instance, McDermott and Roediger (1998) found that participants were more likely to falsely recall words that were not presented but have relevant meanings to words in a list after a delay of one week compared to immediately after the list was presented.

Additionally, in Polišenská et al. (2014) study, participants were presented with a series of sentences that contained implicatures, and they were asked to recall the sentences after a delay of either 10 minutes or 24 hours. The results showed that participants tend to recall the sentences with their implicatures more after the 24-hour delay than after the 10-minute delay.

Moreover, Abadie and Camos (2019) showed that participants could recall the implicatures of the sentences better after the 24-hour delay than the 5-minute delay.

2 Current Study

In our research, we aimed to comprehensively understand the role of prior beliefs on memory recall. We investigated whether individuals' prior beliefs in a just world influence the formulation of false memories during memory reconstruction. Consistent with previous studies on false memories, we designed our experiment around participants' free recalls. As a novel approach, we examined free recalls through two aspects: Accuracy, which refers to how closely the participant's recall matches the original sentence grammatically, and Implicature, which refers to the extent to which the participant relied on pragmatic inferences when recalling the sentence.

Additionally, we developed an algorithm named Implicature Scoring Algorithm (IMS) for automatically assessing participant recalls regarding these reconstructive aspects of memory. This algorithm minimizes human errors and enhances the reproducibility of our findings in future research endeavors. The IMS calculates two main scores: Accuracy Score, which measures how closely the participant's recall matches the original sentence grammatically, and Implicature Score, which assesses how much the participant relied on pragmatic inferences when recalling the sentence.

Our central hypothesis (*Just-World Hypothesis*) was that individuals' prior beliefs in a just world would influence the formulation of false memories. We posited that the contradiction between preexisting beliefs and novel information results in cognitive conflicts and dissonance, leading to psychological discomfort (Festinger, 1957). This discomfort motivates individuals to seek internal consistency via various strategies. The Belief in a Just World Theory suggests that individuals hold prior beliefs about a fair world where people get what they deserve (Lerner, 1980). When novel information violates this belief, it results in cognitive conflicts and dissonance, leading to psychological discomfort. Individuals are motivated to justify their prior beliefs to reduce this discomfort. These justification strategies can lead to reconstructing memories that contradict individuals' beliefs in a fair world (Figure 1). Specifically, we hypothesized that *individuals with strong beliefs in a just world are motivated to reconstruct events, information, or sentences with false memories based on the speaker's implicatures to justify their prior beliefs in situations where those beliefs are violated.*

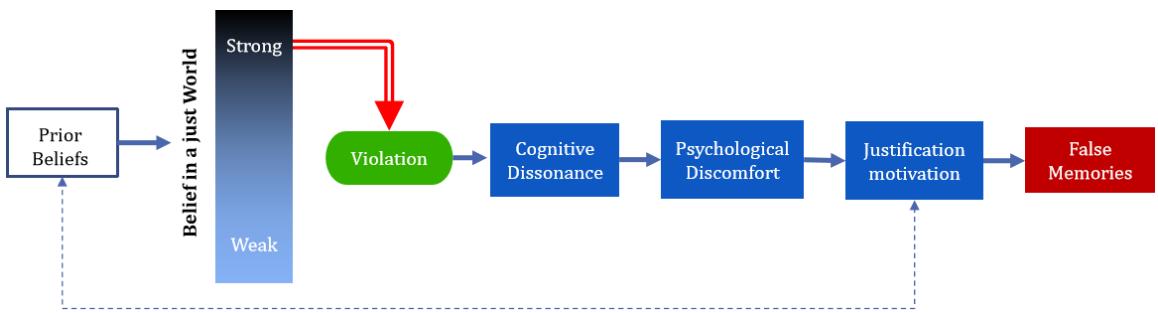


Figure 1 The Graphical Representation of Just World Hypothesis

On the other hand, studies have shown that individuals remember others' negative actions to hold them accountable for future violations (Bell & Buchner, 2012; Mattarozzi et al., 2015; Kroneisen et al., 2015). This tendency can be driven by various reasons, such as maintaining the social group's integrity and avoiding individual threats (Tajfel & Turner, 1979). Meanwhile, the pragmatic inferences of utterances provide more details about the speakers and events (Grice, 1975). Considering those, we hypothesized (*Motivated-by-accountability-goals Hypothesis*) that *listeners would retrieve utterances with their pragmatic inferences based on a speaker's commitment if it is not fulfilled and use them as evidence to hold the speaker accountable for potential future violations.*

Previous research suggests that broken commitments can trigger more robust emotional and cognitive responses, leading to more detailed and accurate memory encoding due to the heightened emotional impact and the need to process the violation (Bonalumi et al., 2019). Respectfully, we included broken and kept commitments in our experimental design to examine how violation of belief in a just world and expectations affect memory recall. On the other hand, previous studies have shown that time gaps between initial exposure and recall impact recall accuracy (Garry et al., 1996; De Gardelle et al., 2012). Smaller time gaps result in better recall accuracy than bigger ones. The bigger time delays between initial exposure and recall weaken memory traces, leading individuals to rely more on generalized schemas or pragmatic inferences rather than specific details (Brainerd & Reyna, 2015; Loftus, 2005).

Incorporating 24 to 72-hour time gaps in our experiment was a deliberate choice to investigate these phenomena under realistic memory conditions. Ebbinghaus's (Murre & Dros, 2015)

findings indicate that significant decay in detail begins within the first 24 hours after learning and continues over the next few days. Specifically, within 24 hours, individuals can forget up to 70% of the information learned, and this percentage increases as more time passes. After 72 hours, the chances of forgetting significantly increase, making recalling memories based on practical assumptions more difficult. This leads to a higher likelihood of complete forgetting (Políšenská et al., 2014; Abadie & Camos, 2019). By introducing a 24 to 72-hour time gap, we ensured that participants' memories would be naturally weakened, making them more susceptible to reconstruction and influenced by their preexisting beliefs.

This approach aligns with the Just World Hypothesis and cognitive dissonance theory. Immediate recall might not fully capture the impact of cognitive dissonance because the emotional and cognitive discomfort caused by the contradiction between belief and reality may not have fully set in. By allowing 24 to 72 hours, we provided enough time for participants to process the dissonance and employ coping strategies, such as reconstructing memories to align with their just-world beliefs. The time gap allows us to observe how participants justify their beliefs over time. The longer duration helps examine whether individuals reconstruct their memories more dramatically to maintain a belief in a just world, thus offering insights into the mechanisms of memory distortion. Furthermore, this time frame simulates a realistic delay, making our findings more applicable to everyday memory recall situations. Emotional responses to violations of just world beliefs (e.g., witnessing injustice) often become more pronounced over time as individuals reflect on the events.

As our experimental design allowed, we included a hypothesis to assess the impact of time delay on memory recall accuracy and implicature formation. More specifically, we hypothesized (*Meaning-to-sentence Hypothesis*) that *individuals tend to reconstruct the utterance based on its pragmatic inference in case of time gaps between the first exposure of the utterance and recall*.

By investigating these hypotheses, we aimed to explore the complex interplay between cognitive and motivational factors in memory recall and provide insights into how prior beliefs shape the reconstruction of memories. This comprehensive approach allowed us to examine how just world beliefs, accountability motivations, and time delays affect memory recall, focusing on forming false memories and the reconstructive processes involved.

3 Method

3.1 Participants

A power analysis conducted using the R package 'pwr2' suggests that a sample size of 600 participants would achieve 84% power with alpha = 0.05, assuming a small to medium effect size of $f^2 = 0.085$. This power analysis was explicitly designed for an ANOVA. We planned to exclude participants who had not completed or returned to the following survey session, failed attention-check questions, and forgot the commitment sentences from the analysis. Our study focused on memory recalls and aimed to investigate potential influences from participants' native languages. We determined that being an adult native English speaker was a selective recruitment criterion and ignored other demographics such as gender and age. Therefore, we planned to recruit 600 native English speakers through Prolific (www.prolific.com) and an additional 150 participants to account for potential exclusions. Our final sample consisted of 750 participants (No Delay conditions: n = 350, Delay condition: n = 400). After data collection, we found that the assumptions for ANOVA were violated, so we conducted non-parametric tests to ensure the robustness and validity of the results. Eventually, 191 participants were recruited but excluded from the analyses due to not completing or returning to the following survey session (n = 56), failing in attention-check questions (n = 1), and forgetting the commitment sentences (n = 141). Consequently, the analysis included 552 participant recalls. The study was approved by the Central European University in Vienna's Psychological Research Ethics Board (PREBO).

3.2 Materials

3.2.1 Vignettes

As part of our experimental design, we created 8 vignettes, each following a consistent structure. The vignettes began with an introduction that established the characters and their situation. It was followed by a dialogue in which the first character either asked for help (in Vignettes 2, 4, and 6), extended an invitation (in Vignettes 1, 3, and 5), posed a question (in Vignette 7), or described a situation (in Vignette 8) to the second character. The second character then

responded with a commitment. The commitment sentences have a direct impact on the recalls. Sentences that do not allow the hearer to reconstruct the speaker's utterances with implicatures could produce biased results. Considering this, we designed the commitment sentences as implied commitment, allowing individuals to construct pragmatic inferences. Finally, the story concluded with an ending that revealed whether the commitment had been fulfilled.

We conducted a norming study to ensure that the participants' recollections were not affected by diverse interpretations of the sentences and to determine whether they could correctly understand the implied meaning of the sentences (i.e., conveying a commitment).

Norming Study – Vignettes

We assessed the reliability of the vignettes by conducting an online survey through the Qualtrics platform. For this purpose, we recruited 40 adults who were native English speakers via Prolific. The survey consisted of eight vignettes, each containing statements designed to test the participants' understanding of implied meanings.

Each vignette comprised three opinion questionnaires and confidence scales. Using a three-option questionnaire (yes, maybe, no), the participants were asked to indicate whether they understood the implied meanings of the statements in the vignettes. For example, one of the vignettes involved a scenario in which Jane responded to Karen's wedding invitation: "I am going to buy a ticket for next Wednesday." The participants were asked whether they thought Jane meant that she would attend Karen's wedding using the three-option questionnaire.

Furthermore, on a seven-point scale, the participants were asked to rate their confidence levels regarding their understanding of the implied meanings. For instance, they were asked to rate their confidence about understanding the implied meaning of the Wedding Vignette on a scale of one to seven, with seven indicating complete confidence.

To determine the reliability of the vignettes, we used a mathematical formula that combined the participants' three-point questionnaire answers and confidence levels (Starmans & Friedman, 2012). We multiplied the participants' three-point questionnaire answers by 1 if their answer was "yes," by -1 if it was "maybe," and by 0 if it was "no." We then multiplied these values by the

participants' confidence ratings and calculated the average of the summed-up results of these multiplications.

$$(1) \quad f(x, y) \begin{cases} y, & x = yes \\ 0, & x = no \\ -y, & x = maybe \end{cases}$$

$$(2) \quad S_{Vignette} = \sum_{i=0} f(x, y)$$

Results

The study found that many participants, at least 73% and above, were able to understand the implied meaning of the target sentences in Vignette 1 Wedding, Vignette 5 Drink, Vignette 6 Draft, Vignette 7 The Lunch, and Vignette 8 The Doggie. On the other hand, participants found the implied meaning of the target sentences in Vignette 2 Concert, Vignette 3 Cinema, and Vignette 4 Plants to be unclear, with a sum of percent of "maybe" and "no" answers being 57% or above.

In addition to the three-point scale results, the confidence scale's findings showed that participants had higher confidence levels (mean 5 and above) in understanding the implied meaning of the target sentences in Vignette 8 Doggie, Vignette 7 The Lunch, Vignette 5 Drink, Vignette 1 Wedding, and Vignette 6 Draft. On the other hand, participants had lower confidence levels (mean 5 or below) in understanding the implied meaning of the target sentences in Vignette 3 Cinema, Vignette 4 Plants, and Vignette 2 Concert compared to the other vignettes.

Furthermore, the reliability scores of the vignettes supported the confidence scale and three-point questionnaire's results, indicating that the implied meanings in Vignette 2 Concert, Vignette 3 Cinema, and Vignette 4 Plants were less understandable compared to the other vignettes by participants. As a result, to prevent the vignette context from influencing the study, Vignette 2 Concert, Vignette 3 Cinema, and Vignette 4 Plants were discarded due to their low-reliability scores (Figure 2).

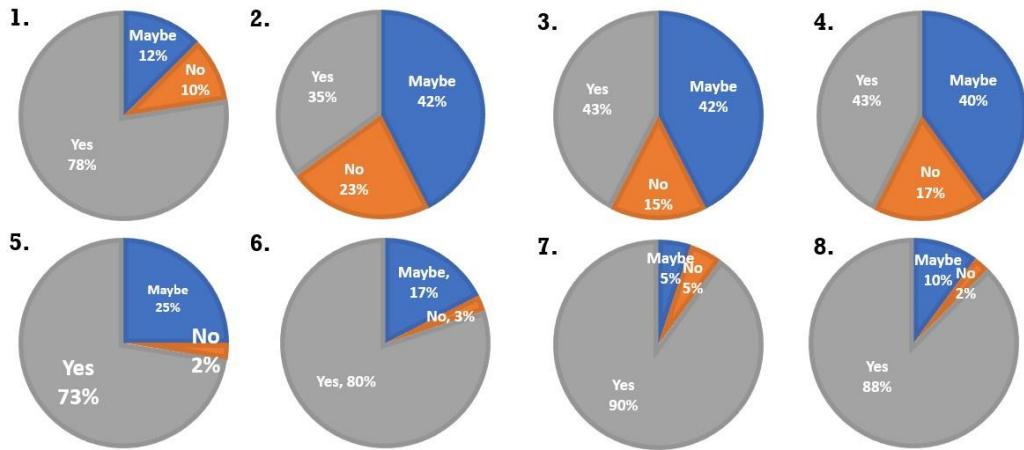


Figure 2 First Norming Study Three-point Questionnaire Answer Distribution

Note. 1. Vignette 1 Wedding (31 yes, 5 maybe, 4 no), 2. Vignette 2 Concert (14 yes, 17 maybe, 9 no), 3. Vignette 3 Cinema (17 yes, 17 maybe, 6 no), 4. Vignette 4 Plants (17 yes, 16 maybe, 7 no), 5. Vignette 5 Drink (29 yes, 10 maybe, 1 no), 6. Vignette 6 Draft (32 yes, 7 maybe, 1 no), 7. Vignette 7 The Lunch (36 yes, 2 maybe, 2 no), 8. Vignette 8 The Doggie (35 yes, 4 maybe, 1 no)

3.2.2 Memory Recall Question

To gather open-ended responses, we designed a set of questions that began with the initial dialogues of the first characters in the vignettes. Participants were then prompted to recall the second characters' responses, which conveyed a sense of commitment. To ensure clarity and accuracy in the responses, participants were explicitly asked to recall the exact words used in the vignettes that conveyed this commitment. We recognized the potential for ambiguity in this question and provided clear instructions to help participants understand the information sought. We aimed to obtain detailed and accurate responses that would provide meaningful insights to our research.

To ensure the accuracy and reliability of our main study, we recognized the possibility of memory recall questions impacting the results. Therefore, we conducted a second norming study with a sample group to verify that our memory recall question was being correctly interpreted. This additional norming study was conducted to eliminate any potential confusion or ambiguity

in the wording of the memory recall question and to provide us with confidence that the main study results would be accurate and valid.

Norming Study – Memory Recall Question

We designed a survey with five vignettes containing statements, each with two different ending conditions: broken and kept commitments. These vignettes were previously tested in a norming study, showing their understandability. We recruited 63 adult native English speakers through Prolific to participate in the survey.

During the survey, we randomly assigned participants to read one of the five vignettes with one of the two ending conditions. We then asked participants to recall the sentences that conveyed the implied commitment at the end of the dialog. For example, in one vignette, Jane replies to Karen's wedding invitation: "I am going to buy a ticket for next Wednesday." We reminded participants of the dialog and asked them to write Jane's exact words.

Finally, we asked participants to rate their confidence in their answers with a seven-point scale. After the survey, we analyzed the participants' memory recalls using the implicature algorithm we developed. The algorithm categorized participants' memory recalls into Implicature and Accuracy scores. It allowed us to evaluate how well participants remembered the commitments made in the vignettes.

Results

The findings showed that five participants scored below 0.3 in both Implicature and Accuracy scores. Upon further scrutiny, these memory recall scores were identified as errors, indicating that the participants did not accurately interpret the memory recall question. However, since the error rate was found to be well within acceptable boundaries, which was approximately 0.06%, it was concluded that the participants had a good understanding of the memory recall question, which implies that the results obtained from the study were reliable and can be used to draw valid conclusions.

3.2.3 Attention Check Question

As part of our research study, we created several vignettes for participants to read and analyze. We formulated simple multiple-choice questions to ensure participants paid close attention to the vignettes and fully comprehended their content. These questions assessed participants' understanding of the vignettes and verified that they had read them thoroughly (Appendix C).

3.2.4 Just World Scale

As per our previous discussion in the literature review section, we assessed the participants' perception of the degree of belief in a just world. For this purpose, we used the Belief in a Just World of Distributive and Procedural Justice For Others. This assessment helped us understand how the participants perceived the world around them regarding fairness and justice, both in the distribution of resources and the procedures employed to distribute them.

3.2.5 Implicature Scoring Algorithm

An accurate scoring system for participants' recalls is essential to ensure a study's reliability and reproducibility. However, manual assessment can be problematic due to scoring limitations and the potential for human error. Therefore, a computational approach was implemented to score recalls in two aspects: Accuracy and Implicature Scores.

- ❖ *Accuracy*: Any discrepancy between the target sentence and recalled sentence (except for typos) regardless of their semantics.
- ❖ *Implicature*: The discrepancy between the target sentence and recalled sentence to the extent that it manifests diagnostic words and the words used in the vignette (and their synonyms), i.e., words that cue the speaker's meaning.

For the Accuracy Score, the first step was removing any punctuation and special symbols from the participant's recall and commitment sentence. The words were then separated into individual units, and a comparison was made between the words in the participant's recall and those in the commitment sentence. Finally, the Accuracy Score was calculated by determining the ratio of the number of words that matched the participant's recall to the number of words in the commitment sentence (Figure 3).

On the other hand, we followed a different computational process for the Implicature Score. As a first step, punctuations and special symbols were removed from the participant's recall and the commitment sentence. In the next step, the words frequently used but do not contribute to the sentence's meaning (stopwords) were removed from the participant's recall and the commitment sentence (Appendix E). The words in the intersection of the remaining words in the participant's recall and the commitment sentence are identified as a set of unmatched words (A). The unmatched words (A) are compared with the diagnostic words. The matched words are identified as Implicatures (I).

Consequently, the Implicature Score was calculated by determining the ratio of matched words (I) to the number of words in the participant's recall (excluding stopwords) (Figure 4).

All the algorithm processes were developed with Python version 3.10.0 and its following packages:

- ❖ Pandas version 2.1.4 (McKinney, 2010): used for accessing, reading, and creating the datasets and Excel files.
- ❖ Regex version 3.8.2 (Van Rossum, 2020): used for removing punctuations and special symbols from sentences.
- ❖ NLTK version 3.8.1 (Bird et al., 2009): used for dividing sentences into individual words (Word Tokenizer).

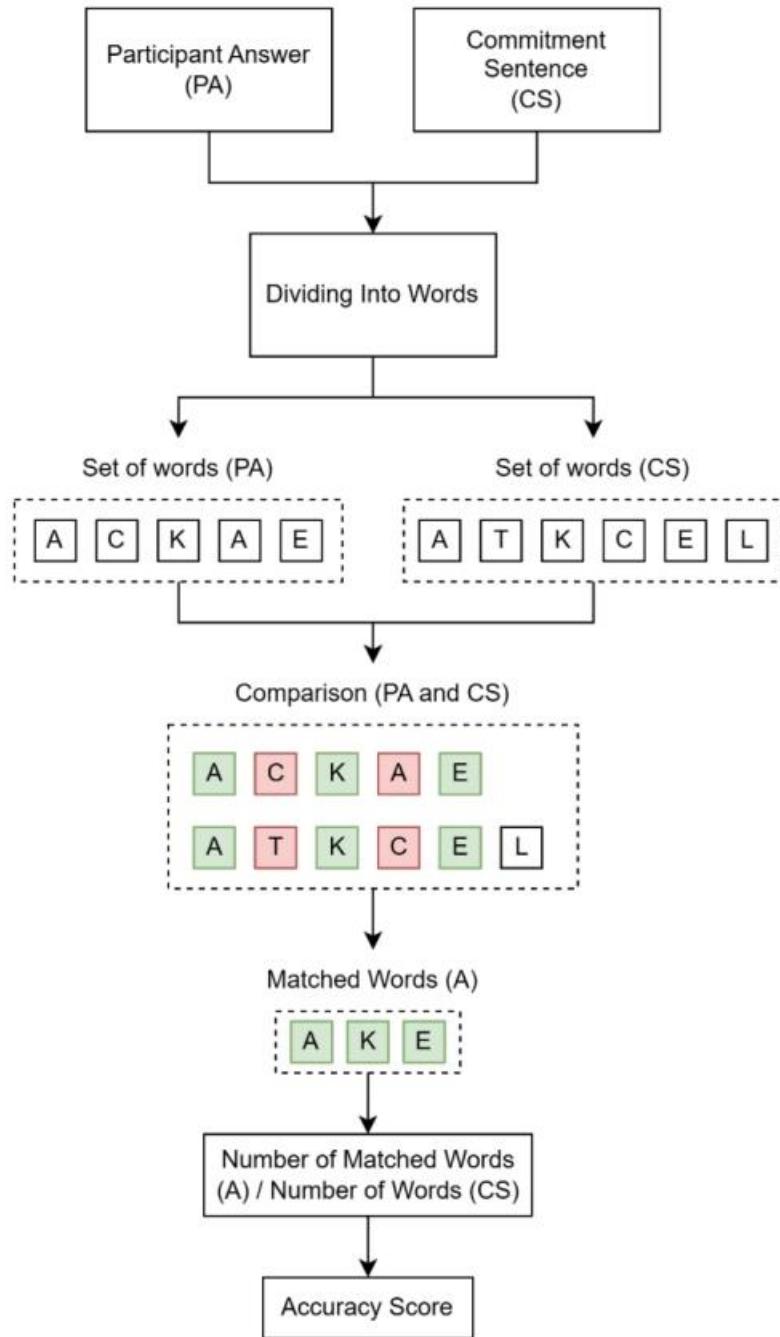


Figure 3 Accuracy Scoring Process Flowchart

Note. The characters such as “A,” “C,” and “K” indicate the words. The colors of the boxes indicate matches (green is a match, and red is not a match).

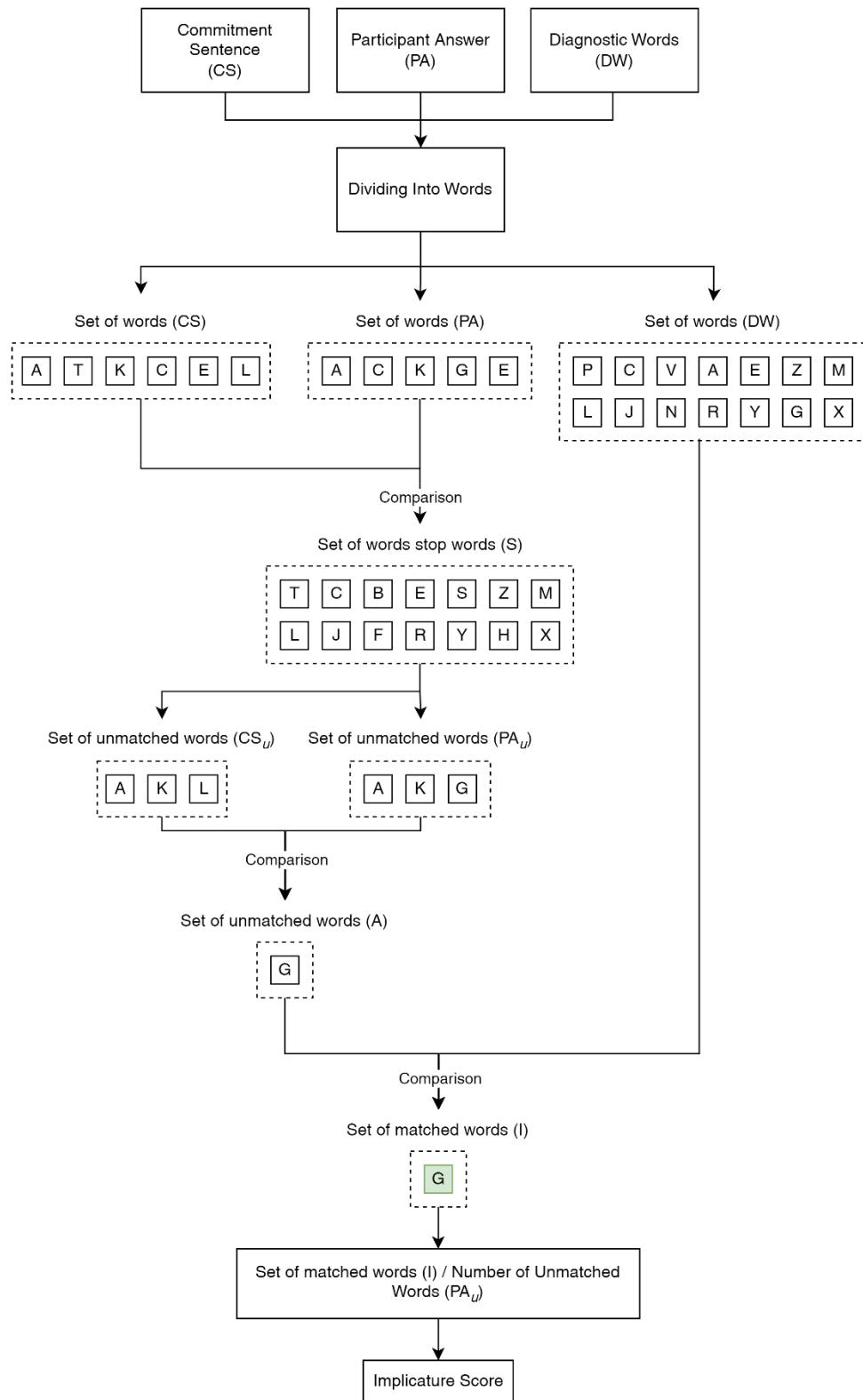


Figure 4 Implicature Scoring Process Flowchart

Diagnostic Words

The individual's recalls usually go beyond what is explicitly stated. Considering this, we implemented the words that connected to the speakers' meaning as a discriminative factor in the Implicature scoring process of the algorithm. We collected the diagnostic words through an experiment we conducted. In the experiment, we recruited 100 adult native English speakers through Prolific, and they forwarded to Qualtrics for the experiment. After the consent form, participants were presented with 5 different vignettes (Wedding, Drink, Draft, The Lunch, The Doggie) and recall questions. We formulated the recall questions to gather the participant's recall, which includes the intention of the speaker's utterance in the vignette. For instance, participants presented the vignette in which Karen mentioned to Jane that she would like to see Jane at her wedding. Then, they were asked to recall what Jane meant when she said, "I am going to buy a ticket for next Wednesday."? Consequently, we removed Stopwords from the gathered participants' recalls, obtaining the diagnostic words for vignettes (Appendix E).

Reliability Test

To assess the reliability of the Implicature Scoring Algorithm, we formulated a reliability test in which five human coders score the participants' recalls in aspects of Implicature. The human coders were recruited in person (two) and online (three). In the reliability test, we introduce the concept of Implicature to coders. After the introduction, coders were presented with an example vignette and participant recalls, including suggested Implicature scores. Then, they were presented with 5 vignettes and 10 participant recalls per vignette. Eventually, they were asked to score the participants' recalls of aspects of Implicature from 0 to 1.

Analysis

We conducted an intraclass correlation analysis to assess the consistency of coders' scores. The Implicature Scores have acceptable internal consistency ($\alpha = .888$).

Furthermore, we followed a specific assessment method to examine the consistency between human coders' scores and the algorithm's scores in Implicature scores. The assessment method follows:

$$\text{Score Difference} = |Algorithm Score - Participant Score|$$

$$\text{Average Difference} = \frac{1}{n} \sum_{i=1}^n |Score Difference_i|$$

$$\text{Reliability Score} = (1 - \text{Average Difference}) \times 100$$

The analysis showed that the Implicature Scoring Algorithm is 67% reliable in assessing the Implicature Scores.

3.3 Experimental Design

We formulated a 2x2 between-subject design, manipulating the time factor at two levels: delay and no delay. Furthermore, we manipulated the commitment factor at two levels: kept and broken. As a result, participants were randomly assigned to one of four conditions: no delay and kept, no delay and broken, delay and kept, and delay and broken (Figure 5).

3.3.1 No Delay Conditions

Before participating in the study, all participants were presented with a consent form (Appendix D). Once consented, they were presented with a vignette, randomly selected from five options, with no ending. The five vignettes were Vignette 8, Doggie; Vignette 7, The Lunch; vignette 5, Drink; vignette 1, Wedding; and Vignette 6, Draft. After reading the vignette, they were asked to answer an attention-check question to ensure they read and understood it.

Next, the participants were presented with an ending for the vignette they had just read. This ending was randomly selected from two possible conditions: broken or kept. The two ending conditions were presented separately to ensure that participants could not see both options and influence their answers.

Finally, at the end of the survey, participants were asked to provide their responses to a memory question, which tested their ability to recall details from the vignette they had read. Additionally, they were asked to rate their overall experience using a belief in a just world scale on a separate page.

3.3.2 Delay Conditions

As in the no-delay condition, the participants were first required to read and approve the consent form. After that, they were presented with the same five vignettes as in the no-delay condition. The vignettes were provided without endings, and the participants were randomly assigned to read one of them. After reading the vignette, they were asked to answer an attention-check question to ensure they read it carefully. At the end of the survey, the participants were informed that they would be invited to participate in the follow-up study by e-mail.

After a one-day delay, participants were invited to the second session of the survey. The email informed them they should participate in the second session no later than two days after receiving it. The participants were again asked to read and approve the consent form in the second session. They were then presented with two separate pages—one for filling in the memory question and the other for filling in the belief in a just world scale. The memory question was designed to assess their recall of the vignette they had read in the previous session, while the just world scale was used to measure their belief in the concept of a just world.

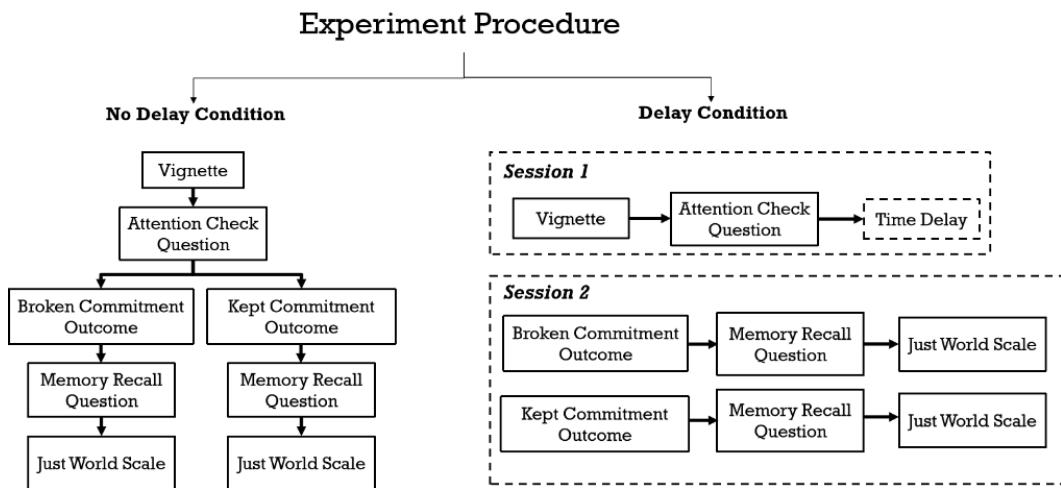


Figure 5 The Flowchart of Experiment Procedure

4 Results

We excluded 25 participants who had not returned for the second session of the delay condition, 31 who left the survey without completing it, and 1 who failed the attention check question from the analysis (Appendix G).

We labeled the participant recalls that scored lower than 0.3 in both the Implicature and Accuracy scores, such as “I do not remember,” “I cannot recall,” and “okay.” as errors. We also excluded the participant recalls they labeled as errors from the analysis (71 delay-broken, 48 delay-kept, 11 no delay-broken, and 11 no delay-kept). The analysis included 552 participant recalls (delay-broken 106, delay-kept 118, no delay-broken 176, no delay-kept 152) (Appendix G). All analyses were conducted using IBM SPSS 25 Statistic software.

Additionally, we conducted a series of chi-square tests (Figure 6) to determine the significance of the delay and commitment conditions on participants’ forgetting (error) rates. The results showed a higher forgetting rate in the delay condition (34.7%) than in no delay condition (6.3%) with a significant Chi-square value ($p = 0.000, df=1$). Furthermore, the forgetting rate was higher in the delay-broken commitment condition (40.1%) than in the delay-kept condition (28.9%) with a significant Chi-square value ($p = .029, df=1$).

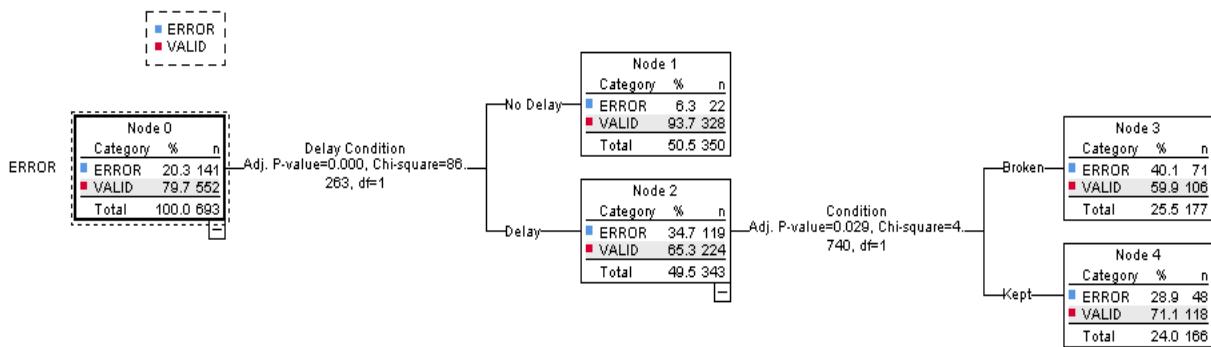


Figure 6 The Tree of Chi-square Tests

We tested the normality of our data with Shapiro-Wilk tests. The results showed that Accuracy and Implicature Scores are not normally distributed in commitment conditions (Table 1) or delay conditions (Table 2). Although we designed and pre-registered the statistical method as

ANOVA, we conducted those normality test results before selecting the statistical methods to test our hypotheses.

Table 1 The Normality Test of Commitment and Delay Conditions

Commitment	Condition	Delay Condition	Kolmogorov-Smirnov			Shapiro-Wilk		
			Statistic	df	Sig.	Statistic	df	Sig.
Broken	Delay	Implicature Score	.300	106	.000	.701	106	.000
		Accuracy Score	.124	106	.000	.949	106	.001
	No	Implicature Score	.519	176	.000	.387	176	.000
		Accuracy Score	.255	176	.000	.826	176	.000
	Kept	Implicature Score	.265	118	.000	.771	118	.000
		Accuracy Score	.131	118	.000	.914	118	.000
Kept	No	Implicature Score	.484	152	.000	.508	152	.000
	Delay	Accuracy Score	.246	152	.000	.834	152	.000

4.1 Just World Hypothesis

Considering the data's non-parametric distribution, we conducted a quantile regression to examine the relationship between Belief in a Just World Distributive Justice for Others score and implicature scores in the commitment conditions. The results indicated that BJW_D was not a significant predictor of implicature scores in the broken commitment condition $\beta=0.00000005$, $p=1$ (Figure 7).

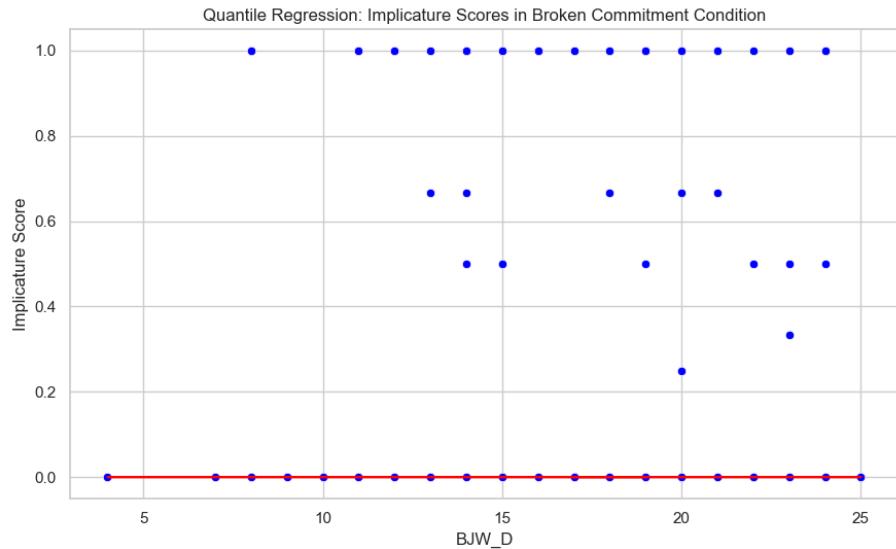


Figure 7 Quantile Regression for Implicature Scores in Broken Commitment

4.2 Motivated-by-accountability-goals Hypothesis

A Mann-Whitney U test was conducted to compare implicature scores between broken and kept commitments (Figure 8). The results indicated that there was a significant difference in Implicature Scores between the Broken (N = 282, Mean Rank = 262.23) and Kept (N = 270, Mean Rank = 291.41) conditions, $U=42095.00$, $W=78680.00$, $Z=2.555$, $p=.011$.

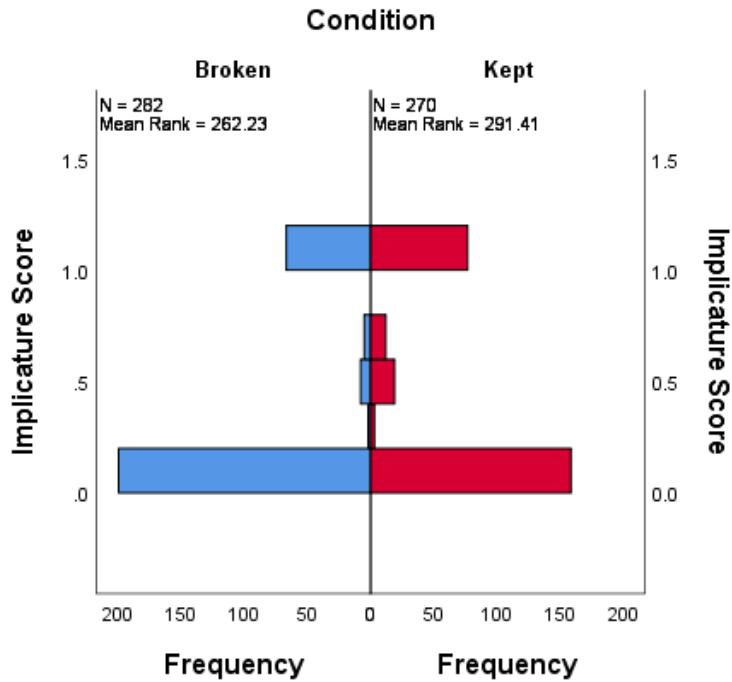


Figure 8 Mann-Whitney U Test for Implicature Scores by Commitment Condition

4.3 Meaning-to-sentence Hypothesis

A Mann-Whitney U test was conducted to compare implicature scores between the delay and no-delay conditions. The results revealed a significant difference in implicature scores between the two conditions, $U = 20647.0, p = .001$, with implicature scores being higher for the delay condition ($N = 224$, Mean Rank = 348.33) than the no-delay condition ($N = 328$, Mean Rank = 277.45) (Figure 9). Additionally, there was a significant difference in accuracy scores between the delay and no-delay conditions (Figure 10), $U = 58261.5, p = .001$, with accuracy scores being higher for the no-delay condition ($N = 328$, Mean Rank = 342.13) than for the delay condition ($N = 224$, Mean Rank = 180.40).

Independent-Samples Mann-Whitney U Test

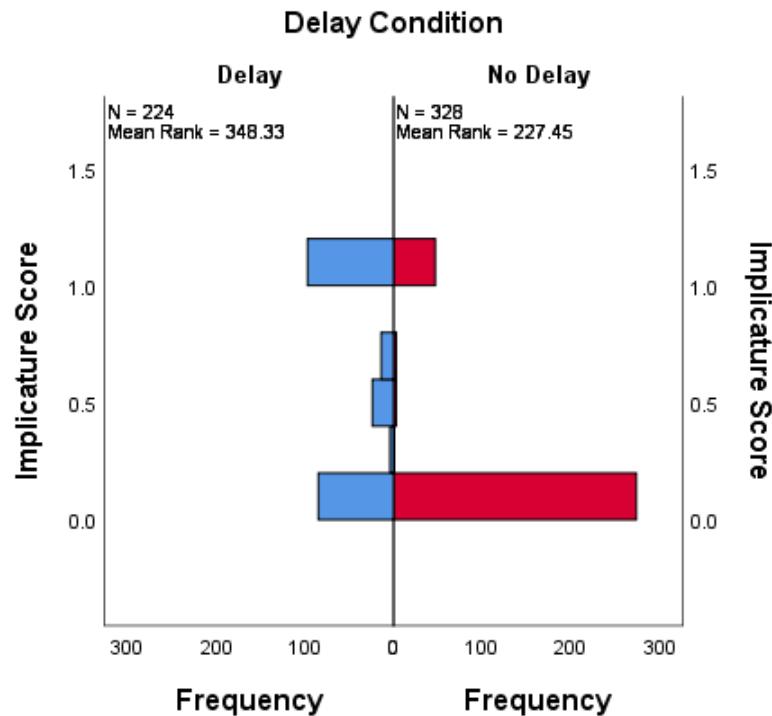


Figure 9 Mann-Whitney U Test for Implicature Scores by Delay Condition

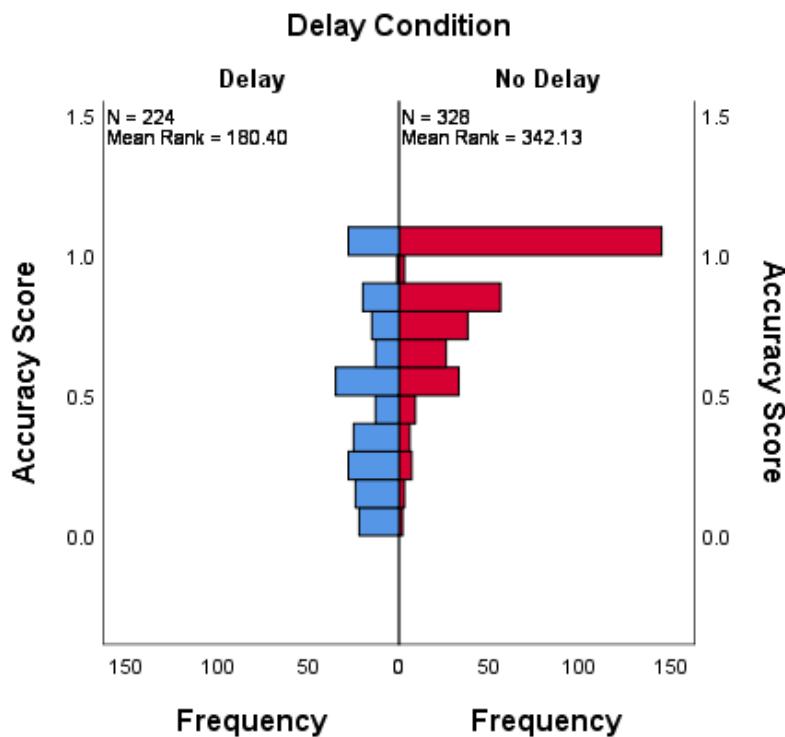


Figure 10 Mann-Whitney U Test for Accuracy Scores by Delay Condition

4.4 Other Remarks

We conducted an explanatory quantile regression to the relationship between Belief in a Just World Distributive Justice for Others score and implicature scores in the kept commitment condition. The results showed that BJW_D was not a significant predictor of implicature scores in the broken commitment condition $\beta = 0.00000003, p = 1$ (Figure 11).

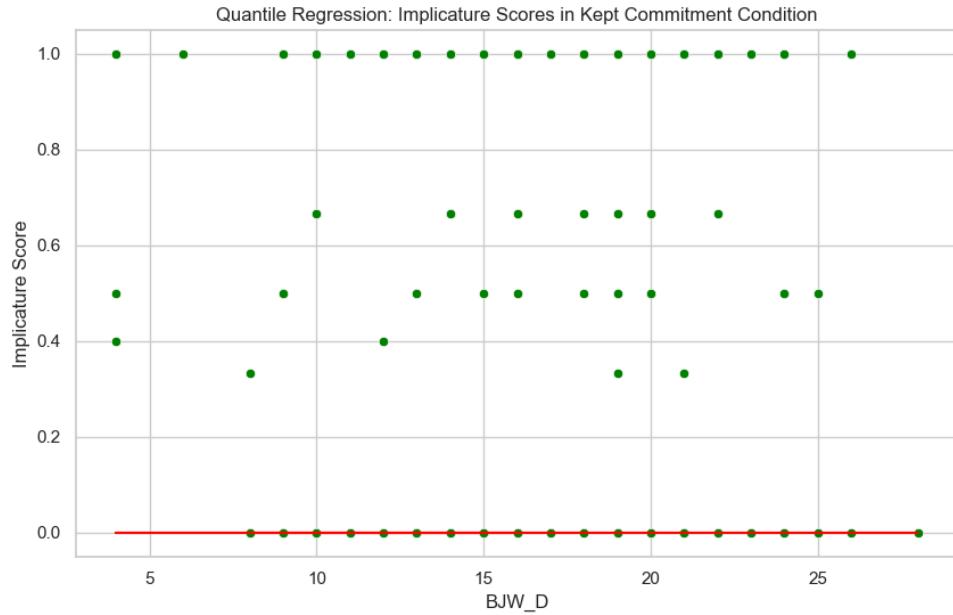


Figure 11 Quantile Regression for Implicature Scores in Kept Commitment

We conducted an exploratory decision tree analysis (Figure 12) to reveal the significant effects of delay condition and commitment outcome on implicature scores. The root node (Node 0) represents the entire sample ($N = 552$) with a mean implicature score of 0.308 ($SD = 0.439$). Participants in the no delay condition (Node 1, $n = 328, M = 0.155, SD = 0.355$) and delay condition (Node 2, $n = 224, M = 0.533, SD = 0.453$) differ significantly (Adj. p -value = 0.000, $F = 120.041, df1 = 1, df2 = 550$). Within the no-delay group, those who were assigned to kept commitment (Node 3, $n = 152, M = 0.200, SD = 0.393$) had higher mean implicature scores compared to those who were assigned to broken commitment (Node 4, $n = 176, M = 0.116, SD = 0.315$) with an Adj. p -value of 0.035 ($F = 4.504, df1 = 1, df2 = 326$). This suggests that both

the delay condition and the nature of the commitment outcome significantly influence the implicature score, with delays and kept commitments associated with higher implicature scores.

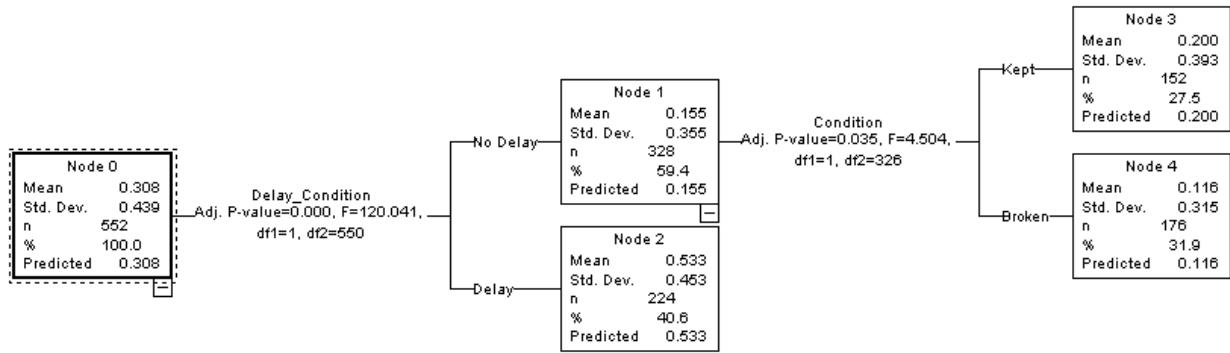


Figure 12 Decision Tree of Delay and Commitment Conditions on Implicature

A quantile regression examined the relationship between Belief in a Just World (Distributive Justice) and Accuracy Scores in the commitment conditions. The results indicated that BJW_D was not a significant predictor of accuracy scores in the neither broken (Figure 13) $\beta = -0.0062$, $p = 0.256$ nor kept commitment conditions (Figure 14) $\beta = 0.0062$, $p = 0.451$.

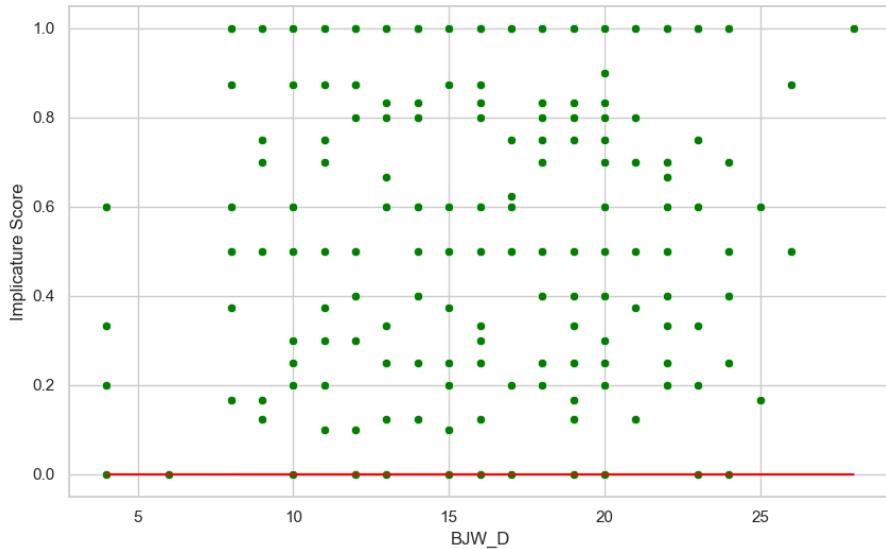


Figure 13 Quantile Regression for Accuracy Scores in Broken Commitment

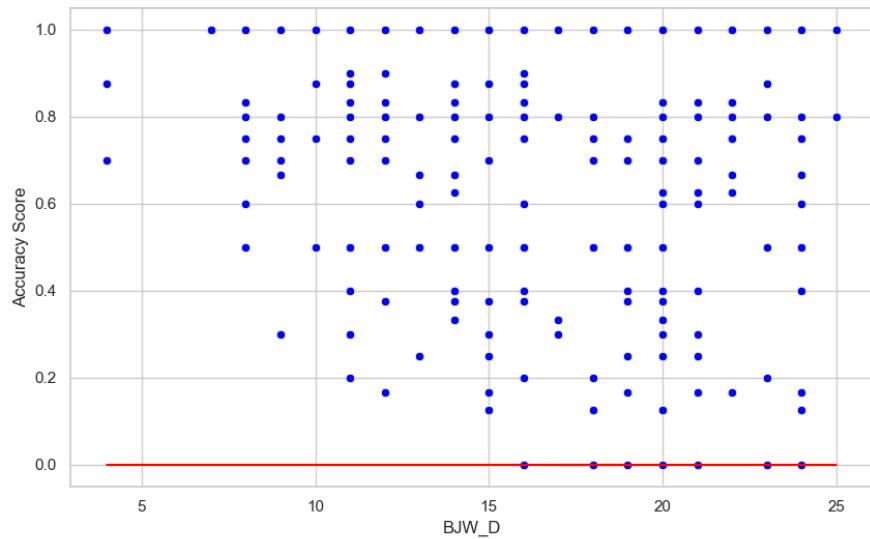


Figure 14 Quantile Regression for Accuracy Scores in Kept Commitment

We conducted an exploratory Mann-Whitney U test to compare accuracy scores between broken and kept commitments (Figure 15). The results indicated a significant difference in accuracy scores between broken and kept commitments: $U=33262.5, p=.009$. Specifically, accuracy scores were higher for the broken commitment condition than for the kept commitment condition.

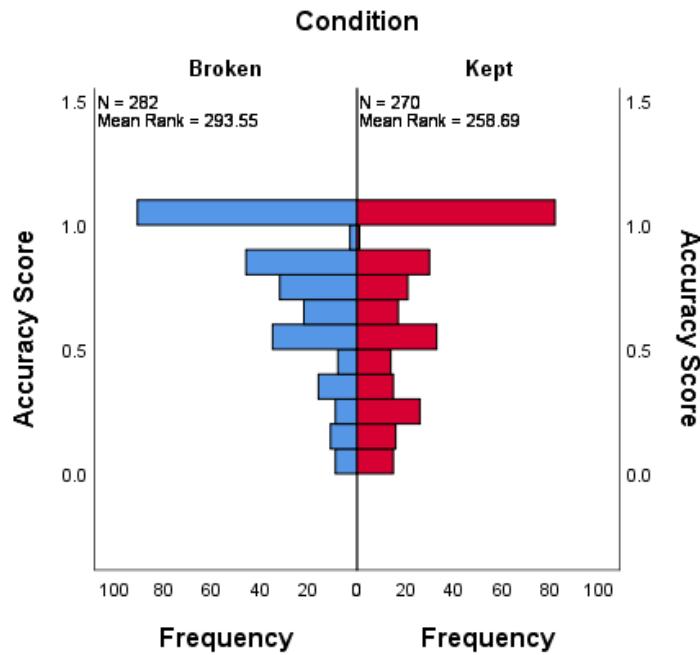


Figure 15 Mann-Whitney U Test for Accuracy Scores by Commitment Condition

5 Discussion

5.1 Forgetting

Based on the multistore memory model, information stored in long-term memory tends to fade over time if it is not reinforced by new information or repeated frequently. The time gap between the initial exposure to information and its recall significantly influences the degree of forgetting. Our study supported the suggestions of the memory model. The results showed that participants who were assigned a 24—to 48-hour delay between the first exposure and recall of written content experienced significant forgetting.

Furthermore, our results revealed that participants assigned both the delay time condition and the broken commitment condition were likelier to forget the commitment sentences than those assigned the delay-kept commitment condition. Our study defined the broken commitment condition as an outcome of the second character's unfulfilled commitment. From an emotional perspective, such unfulfilled outcomes might have negatively affected the participants' emotional states.

Individuals tend to remember negative memories in a more positive light due to a phenomenon called fading affect bias (Markus & Wurf, 1987; Pyszczynski & Greenberg, 1987; Skowronski et al., 1991; Walker & Skowronski, 2009). This optimistic view of negative memories helps people maintain their positive self-image. Combining this bias and the multi-store memory model can explain the relationship between forgetting and a delay-broken condition. When individuals are in a negative emotional state due to an unsatisfied commitment, the time delay can worsen the information loss in long-term memory, leading to forgetting the commitment sentence. Thus, the negative emotional state created by the unsatisfied commitment can make the information loosened by time delay even more susceptible to forgetting.

5.2 Implicature Scoring Algorithm

Implicatures are pragmatic inferences arising from the speaker's utterance. Hearers reconstruct the speaker's utterances with pragmatic inferences based on context. Individual differences such as prior knowledge, experiences, and cultural background are critical to reconstructing

pragmatic inferences, making implicatures a complex and challenging aspect of language to analyze.

In our study, we developed a computational model to assess the extent of implicatures individuals resort to in retrieval. Unlike machine learning models, we built our model based on pre-defined rules. It allowed us to develop and utilize a model without data gathering or training. On the other hand, the rule-based model approach is less flexible than a machine learning model due to various factors that affect the implicature reconstruction. The experimental data showed that participants with the same vignette recalled the speaker's utterance with its implicatures in various grammatical structures and word choices. For instance, in the memory recall task of the Wedding vignette, participants resorted to numerous implicatures, such as "I would love to come but unfortunately I am busy with work," "SEE YOU ON WEDNESDAY," "I'll be happy to."

Consequently, our computational model scored the extent of implicature in participant recalls with 67% accuracy compared to the human coders' scores. Interestingly, our accuracy score is similar to previous computational attempts to calculate implicature. For instance, Li's (2022) study shows 54-64% accuracy in contextual utterances, while Zheng et al.'s (2021) LSTM-Gen Model shows 62-77% in implicature recovery tasks.

5.3 Just World Hypothesis

In the Just World Hypothesis, we hypothesized that individuals with strong beliefs in a just world are motivated to maintain their belief in a just world with false memories in case of violation of them. Regarding this, we predicted that we would observe a significant relation between Belief in a Just World Distributive Justice for Others scores and implicature scores in the broken commitment condition. However, the quantile regression analysis results did not show significant relationships between Belief in a Just World Distributive Justice for Others scores (BJW_D) and implicature scores in broken or kept commitment conditions. Additionally, we conducted an exploratory quantile regression analysis between accuracy scores and Belief in a Just World Distributive Justice for Others scores (BJW_D) for broken and kept commitment conditions. Consistent with the implicature scores, results did not show a significant relationship between Belief in a Just World Distributive Justice for Others (BJW_D) and accuracy scores in

broken or kept commitment conditions. This finding is critical and can be better understood through the various concepts and theories discussed in the literature review section.

According to Cognitive Dissonance Theory (Festinger, 1957), individuals experience psychological discomfort when their beliefs are contradicted by new information. The Belief in a Just World (BJW) theory suggests that people are motivated to see the world as fair, where individuals get what they deserve (Lerner, 1980). When faced with information that contradicts this belief, such as a broken commitment, individuals might experience cognitive dissonance. They could reconstruct their memories to align with their just world beliefs to reduce this discomfort.

However, our results indicate that this motivation does not significantly affect the formation of false memories regarding implicature and accuracy. This suggests that the emotional buffering effect of BJW, which allows individuals to experience less distress and more positive emotions when faced with injustice (Hafer & Bègue, 2005), might not extend to the formation of false memories in our experimental conditions. This aligns with the idea that the context and emotional intensity required to trigger such cognitive biases were possibly not sufficiently strong in our vignettes. Supporting this idea, previous studies (Gerrig & Rapp, 2004; Green & Brock, 2000) proved that strong character traits and emotionally charged scenarios are crucial for eliciting robust cognitive and emotional responses, which our vignettes may have lacked.

Furthermore, according to Schema Theory, information that contradicts pre-existing schemas might be altered or forgotten to fit the schema. From this view, BJW was expected to act as a schema, leading individuals to reconstruct memories that support their belief in a fair world. Similar to cognitive dissonance, the reason behind our results could be the lack of emotionally impactful content in the vignettes, which might have been necessary to activate BJW-related schemas robustly.

The Source-Monitoring Framework explains false memories as errors in determining the origin of a memory (Johnson et al., 1993). From this view, individuals with strong BJW were expected to misattribute information that violates their just-world belief, thus forming false memories.

The absence of a significant relationship in our results suggests that BJW might not strongly influence source-monitoring errors in our experimental context. This could be due to the

complexity of experimentally manipulating BJW and the challenges in creating scenarios that accurately reflect real-world violations of this belief (Hafer & Bègue, 2005).

Additionally, the Fuzzy-Trace Theory suggests that individuals rely more on gist traces than verbatim ones, which can lead to distortions (Brainerd & Reyna, 2005). Regarding our experimental design, individuals with strong BJW were expected to rely on gist traces that align with their belief in a fair world, resulting in false memories.

However, our results indicated that participants did not rely on BJW-related gist traces to reconstruct memories. This could be because the scenarios in our vignettes were not sufficiently salient or emotionally charged to activate such processes, which is essential for activating gist (Tun et al., 1998).

The Dual-Processing Model suggests that memory processing occurs through two systems: an automatic, fast, and unconscious system (System 1) and a controlled, slow, and conscious system (System 2) (Gawronski & Creighton, 2013). From this view, it was expected that System 1 would be influenced by BJW to quickly generate justifications or distortions, which System 2 might not adequately correct.

Our results showed that System 1's automatic processes were not sufficiently triggered to affect memory reconstruction in the context of our study. This aligns with findings by Gallo et al. (2001), which suggest that System 1's influence on memory distortion is limited without strong initial impressions or emotional triggers.

The Self-Memory System Model suggests that individuals prioritize memories relevant to their self-concept (Conway & Pleydell-Pearce, 2000). Regarding this, BJW was expected to influence memory recall by aligning memories with the self-concept of a just world.

However, our findings indicate that BJW did not significantly affect memory recall in our experimental conditions. This could be because the vignettes did not strongly relate to participants' self-concepts or because the situations were not personally relevant enough to trigger self-memory processes. Research by Mazzoni and Memon (2003) supports the idea that personal relevance is critical in influencing memory through self-concept.

Consequently, our study's findings highlight the need for more nuanced experimental designs to better differentiate between personal and general BJW. Future research should ensure that

vignettes' emotional and perceived injustice components are strong enough to trigger BJW-related cognitive processes. Additionally, controlling for individual differences in BJW, possibly influenced by personal experiences and cultural backgrounds, can provide a more accurate understanding of how BJW influences memory formation and false memory occurrence.

By addressing these challenges, future research can delve deeper into how BJW and other cognitive and motivational factors interact to shape memory processes. Understanding these dynamics is crucial for comprehending how individuals maintain their beliefs in the face of contradictory information and how this impacts their recollection and interpretation of past events.

5.4 Motivated-by-accountability-goals Hypothesis

The results showed a significant difference in implicature scores between broken and kept commitment conditions. However, in contrast to our predictions, implicature scores were higher in the kept commitment condition than the broken commitment condition. Moreover, the accuracy scores were higher in the kept commitment condition than in the broken commitment condition. Those results show that individuals tended to remember the speaker's utterance more precisely when the implied commitment was unsatisfied.

However, various factors discussed in the literature could explain the unexpected results. Firstly, relevance theory might explain why kept commitments might lead to higher implicature scores. According to relevance theory, individuals prioritize information that has significant cognitive benefits with minimal effort (Sperber & Wilson, 1995, 2002). Kept commitments could be perceived as more relevant than broken ones as they meet expectations and confirm the communicator's reliability (Bonalumi et al., 2020). On the other hand, implicatures of the speaker's utterances provide more cues about the event and the speaker, providing individuals with cognitive benefits and minimal effort to retain memory, which can result in higher implicature scores in kept than broken commitment conditions.

The higher accuracy scores in the broken commitment condition might initially seem counterintuitive. However, one possible explanation might be the emotional and cognitive impact of violated expectations. Broken commitments often trigger strong emotional responses such as disappointment or resentment. These negative emotions could increase the intensity of

individuals' cognitive processing to understand the violation and its implications (Bonalumi et al., 2019). This increased cognitive effort can result in more detailed and accurate memory encoding.

Additionally, normative judgments about broken commitments might play a role. When commitments are broken, individuals tend to make negative normative judgments about the communicator (Bonalumi et al., 2019). These judgments can lead to a more analytical and detailed processing of the event as people seek to justify their negative evaluations. This heightened analytical processing can enhance memory accuracy for the details of the broken commitment rather than motivate inaccurate recalls.

Lastly, the trust dynamics might further explain the findings. While kept commitments reinforce trust, broken commitments can disrupt it. The disruption of trust can lead individuals to scrutinize the details of the broken commitment more closely, again enhancing memory accuracy. The breach of trust may make the event more memorable as individuals are more likely to remember events that violate their expectations and social norms (Tajfel & Turner, 1979).

5.5 Meaning-to-sentence Hypothesis

The significant difference in implicature scores between the two conditions suggests that a delay notably impacts the interpretation of implicatures. Specifically, implicature scores were higher in the delay condition, indicating that the delay might lead to a more thorough consideration of the implicatures in the given sentences. Consistent with the implicature scores, the significant difference in accuracy scores between delay and no-delay conditions indicates that the absence of a delay is associated with higher accuracy in processing sentences. More specifically, the time delay between exposure and recall is one of the discriminative factors in recall accuracy and implicature formulation.

Additionally, those results support various theories of memory. For instance, the Temporal Context Model (TCM) suggests that memories are encoded within an evolving temporal context. When a delay occurs between encoding and retrieval, the temporal context at retrieval differs significantly from the encoding context. That discrepancy between temporal and encoding context leads to a reliance on generalized schemas or implicatures to reconstruct the memory

(Moscovitch & Craik, 1976; Anderson, 1981; Kumaran et al., 2009). Consequently, participants in the delay condition might have relied more on general knowledge structures than specific details of the original event, resulting in higher implicature scores (Bransford & Johnson, 1972; Alba & Hasher, 1983). In contrast, without delay, the encoding and retrieval contexts are more similar, resulting in higher accuracy scores (Kumaran et al., 2009).

Similarly, the Contextual Binding Theory suggests that memories are formed by binding different elements (e.g., sensory inputs, contextual details) into a cohesive memory trace. If not reinforced, the strength of these bindings can weaken over time. Therefore, the time delay might cause participants to lose specific contextual details, leading them to reconstruct memories based on remaining, more generalized elements, thus increasing implicature use (De Gardelle et al., 2012; Garry et al., 1996). On the other hand, when there is no delay, these bindings remain strong. It allows for more accurate recall of specific details (Brainerd & Reyna, 2015), resulting in higher accuracy scores.

5.5 Limitations

In developing our implicature scoring algorithm, we aimed to create a robust automatic assessment method to minimize the human errors inherent in manual assessment. Our approach was grounded in fundamental linguistic principles to evaluate participants' recall accuracy and implicature. While we endeavored to ensure the reliability and accuracy of our algorithm through rigorous testing, several limitations must be acknowledged.

We tested the reliability of our algorithm by comparing its scores with those provided by human coders for randomly selected recalls. This approach was intended to validate the algorithm's performance against a human benchmark. However, the task-specific nature of our algorithm posed a unique challenge. The accuracy of this validation method hinged on the human coders' understanding of the implicature concept. This reliance introduces a paradox: We aimed to reduce human error through automation, yet we validated our tool against potentially flawed human judgments. This dependency on human coders' interpretations underscores a fundamental limitation of our study.

Experimentally manipulating Belief in a Just World (BJW) in a laboratory proved challenging. Real-world scenarios where BJW operates involve complex and multifaceted experiences that

are difficult to replicate experimentally. Our vignettes may not have captured the full emotional and contextual richness required to elicit strong BJW-related responses. This limitation highlights the need for more sophisticated and ecologically valid experimental designs in future research.

The emotional impact of our vignettes and the presence of notable personality traits in the protagonists were possibly insufficient to trigger BJW-related cognitive biases. Previous research indicates that characters with distinct traits elicit more robust emotional and cognitive responses, influencing memory encoding and recall (Gerrig & Rapp, 2004; Green & Brock, 2000). Our vignettes' relatively neutral or ambiguous character traits might have reduced the emotional engagement necessary to activate BJW-related processes.

Individuals vary widely in their inherent BJW, influenced by personal experiences, cultural backgrounds, and socialization processes. Our study may not have adequately controlled for these individual differences, potentially impacting the generalizability of our findings. Future research should incorporate more nuanced measures to account for personal and cultural variations in BJW.

Our study focused on implicature and accuracy scores as measures of false memory formation. While these metrics provide valuable insights, they represent only a subset of the broader cognitive and motivational factors involved in memory processes. Future studies should explore additional dimensions of memory reconstruction, such as the emotional valence of memories, the role of suggestibility, and the impact of social and contextual cues.

Although our study was grounded in cognitive and linguistic theories, we did not directly measure the neurobiological underpinnings of memory formation and recall. Advances in neuroimaging techniques offer opportunities to explore how brain regions such as the hippocampus, prefrontal cortex, and amygdala contribute to false memory formation. Incorporating neurobiological measures could provide a more comprehensive understanding of the mechanisms underlying BJW-related memory processes.

The task-specific nature of our algorithm means it was designed with types of implicature and recall in mind. While this specificity enhances its reliability within the study's context, it may limit its applicability to other types of memory tasks or different linguistic contexts.

Generalizing our findings to broader contexts will require further refinement and validation of the algorithm across diverse memory and recall scenarios.

While comparing the algorithm's output to human coders' scores was necessary for validation, this inherently tied the algorithm's accuracy to human judgment's potential biases and limitations. Future work should aim to refine the algorithm further and explore alternative validation methods that reduce dependency on human interpretation, perhaps using more extensive and varied datasets or advanced machine learning techniques.

In conclusion, while our implicature scoring algorithm represents a significant step forward in reducing human error in memory assessment, these limitations underscore the complexity of memory processes and the need for continued methodological advancements. Addressing these limitations in future research will enhance our understanding of how BJW and other cognitive biases influence memory formation and contribute to the broader field of memory studies.

Conclusion

This study explored the impact of prior beliefs on memory recall and the reconstructive processes involved, focusing on how the belief in a just world (BJW) affects false memory formation during memory recall. We hypothesized that individuals strongly believing in a fair world would be more inclined to form false memories, especially when their beliefs were violated. Additionally, we posited that these individuals would reconstruct their memories using implicatures when commitments were not fulfilled, and there was a delay between exposure and recall.

We developed an innovative algorithm to automatically assess discrepancies between recall and commitment sentences regarding accuracy and implicature to test these hypotheses. This algorithm reduces the likelihood of human error by analyzing grammatical distinctions to measure accuracy and examining pragmatic inferences to measure implicature. Participants were presented with vignettes and asked to recall commitment sentences under various conditions, including broken versus kept commitments and delayed versus immediate recall.

The results of our experiment did not support the initial hypotheses regarding BJW and false memory formation. Instead, they highlighted the significant roles of time delay and the nature of commitments in shaping memory recall. Participants who experienced a time delay exhibited higher implicature scores and lower accuracy, indicating that time influences the reconstruction of memories towards more pragmatic inferences. Interestingly, participants recalled broken commitments more accurately than kept ones, likely due to the heightened emotional and cognitive impact of violated expectations.

These findings underscore the complexity of memory processes and suggest that factors such as emotional salience and the nature of commitments significantly influence memory reconstruction. The development of our automatic assessment method for memory recall represents a significant advancement, offering a valuable tool for future research to enhance the reliability and reproducibility of findings in memory studies. Further research is needed to continue disentangling the various cognitive and motivational factors that influence memory formation and recall, providing deeper insights into how beliefs and biases shape our recollections.

Moreover, our study opens several avenues for future research. One potential direction is to investigate the influence of other cognitive biases on memory recall and false memory formation. For example, exploring how confirmation or hindsight bias might interact with BJW could provide a more comprehensive understanding of the interplay between cognitive biases and memory processes. Additionally, longitudinal studies could examine how the stability or change in BJW over time affects memory recall, offering insights into the dynamic nature of belief systems and their impact on memory.

Another promising area for future research is the application of our implicature scoring algorithm in different contexts. For instance, its use in clinical settings to assess memory accuracy in patients with memory-related disorders could provide valuable diagnostic and therapeutic insights. Furthermore, integrating this algorithm with neuroimaging techniques could help identify the neural correlates of memory reconstruction processes, shedding light on the brain mechanisms underlying the influence of BJW on memory.

In conclusion, while our implicature scoring algorithm marks a step forward in reducing human error in memory assessment, the study's limitations highlight the need for ongoing methodological improvements. Addressing these limitations will enhance our understanding of how cognitive biases like BJW influence memory processes, contributing to the broader field of memory studies. By continuing to refine our methodologies and explore new research directions, we can further unravel the intricate relationship between beliefs, memory, and cognition, ultimately enhancing our ability to understand and mitigate the effects of cognitive biases on memory.

References

Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, 148(8), 1312–1334.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93(2), 203–231. <https://doi.org/10.1037/0033-2909.93.2.203>

American Psychiatric Association, DSM-5 Task Force. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5™* (5th ed.). American Psychiatric Publishing, Inc.. <https://doi.org/10.1176/appi.books.9780890425596>

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. Spence (Ed.), *The psychology of learning and motivation* (Vol. 2). Oxford, England: Academic Press.

Baddeley, A. (2006). Working Memory. *Working Memory and Education*, 1–31. doi:10.1016/b978-012554465-8/50003-x

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, 61(3), 438–456.

Baddeley, A., & Hitch, G. J. (2007). Working memory. In Oxford University Press eBooks (pp. 1–20). <https://doi.org/10.1093/acprof:oso/9780198570394.003.0001>

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology*, 91(4), 612.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Bell, R., & Buchner, A. (2012). How Adaptive Is Memory for Cheaters? *Current Directions in Psychological Science*, 21(6), 403–408. <https://doi.org/10.1177/0963721412458525>

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Bonalumi, F., Isella, M., & Michael, J. (2019). Cueing implicit commitment. *Review of Philosophy and Psychology*, 10(4), 669–688.

Bonalumi, F., Scott-Phillips, T., Tacha, J., & Heintz, C. (2020). Commitment and communication: Are we committed to what we mean, or what we say?. *Language and Cognition*, 12(2), 360-384.

Brainerd, C. J., & Reyna, V. F. (2005). The science of false memory. In Oxford University Press eBooks. <https://doi.org/10.1093/acprof:oso/9780195154054.001.0001>

Braisby, N., & Gellatly, A. (2012). *Cognitive psychology*. Oxford University Press.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 717-726. [https://doi.org/10.1016/S0022-5371\(72\)80006-9](https://doi.org/10.1016/S0022-5371(72)80006-9)

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5(6), 673–678. <https://doi.org/10.3758/BF03197414>

Brewer, W. F., & Lichtenstein, E. H. (1975). Recall of logical and pragmatic implications in sentences with dichotomous and continuous antonyms. *Memory & Cognition*, 3, 315-318.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. Center for the Study of Reading Technical Report; no. 325.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive psychology*, 13(2), 207-230.

Brewin, C. R. (2007). Autobiographical memory and the self in a clinical perspective. *Memory*, 15(4), 405-410.

Brown, A. L., & Smiley, S. S. (1977). Rating the importance of structural units of prose passage. *Bulletin*, 114, 3–28.

Callan, M. J., Kay, A. C., Davidenko, N., & Ellard, J. H. (2009). The effects of justice motivation on memory for self-and other-relevant events. *Journal of Experimental Social Psychology*, 45(4), 614-623.

Callan, M. J., Powell, N. G., & Ellard, J. H. (2007). The consequences of victim physical attractiveness on reactions to injustice: The role of observers' belief in a just world. *Social Justice Research*, 20, 433-456.

Cherry, C. (1966). On human communication.

Christianson, S. Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion*, 5(2), 81-108.

Colombo, M., Hayne, H., & Rovee-Collier, C. (2000). The Development of Implicit and Explicit Memory. *The Development of Implicit and Explicit Memory*, 1-334.

Colosio, M., Shestakova, A., Nikulin, V. V., Blagovechtchenski, E., & Klucharev, V. (2017). Neural Mechanisms of Cognitive Dissonance (Revised): An EEG Study. *The Journal of Neuroscience*, 37(20), 5074–5083. doi:10.1523/jneurosci.3209-16.2017

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological review*, 107(2), 261-288.

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social cognition*, 22(5), 491-529.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind: Evolutionary psychology and the generation of culture*, 163, 163-228.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. Anderson, J. R. (1981). Effects of prior knowledge on memory. *Psychological Review*, 88(4), 368-394. <https://doi.org/10.1037/0033-295X.88.4.368>

Dalbert, C. (1998). Belief in a Just World, Well-Being, and Coping with an Unjust Fate. In: Montada, L., Lerner, M.J. (eds) *Responses to Victimization and Belief in a Just World. Critical Issues in Social Justice*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-6418-5_6

Dalbert, C., & Sallay, H. (2004). The justice motive in adolescence and young adulthood. Routledge.

Dalton, A. N., & Huang, L. (2014). Motivated Forgetting in Response to Social Identity Threat. *Journal of Consumer Research*, 40(6), 1017–1038. doi:10.1086/674198

Dawtry, R. J., Cozzolino, P. J., & Callan, M. J. (2019). I blame therefore it was: The effects of blaming victims of injustice on superstitious thinking. *Personality and Social Psychology Bulletin*, 45(3), 391-404.

De Gardelle, V., & Summerfield, C. (2012). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences*, 109(18), 6827-6832. <https://doi.org/10.1073/pnas.1110318109>

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17-22.

Dennis, N. A., Kim, H., & Cabeza, R. (2007). Effects of aging on true and false memory formation: An fMRI study. *Neuropsychologia*, 45(14), 3157-3166.

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep medicine reviews*, 13(5), 309-321.

Etam, B., Miele, D. B., & Higgins, E. T. (2013). Motivated remembering: Remembering as accessibility and accessibility as motivational relevance. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 463–475). Oxford University Press.Erlbaum.

Fazio, L. K., & Marsh, E. J. (2010). Correcting false memories. *Psychological science*, 21(6), 801-803.

Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press.

Festinger, L. (1968). A theory of cognitive dissonance. Stanford University Press.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288–299. <https://doi.org/10.1037/0096-1523.1.3.288>

Freud, S. (1953). On aphasia; a critical study.

Frink, D. & Klimoski, R. (2004). Advancing accountability theory and practice: Introduction to the human resource management review special edition. *Human Resource Management Review*. 14. 1-17. 10.1016/j.hrmr.2004.02.001.

Furnham, A., & Procter, E. (1989). Belief in a just world: Review and critique of the individual difference literature. *British journal of social psychology*, 28(4), 365-384.

Gallo, D. A., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, 8, 579-586.

Gardiner, J. M., & Java, R. I. (1993). Recognising and remembering. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 163–188). Lawrence Erlbaum Associates, Inc.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, 3(2), 208-214. <https://doi.org/10.3758/BF03212420>

Gawronski, B., & Creighton, L. A. (2013). Dual process theories 14. *The Oxford handbook of social cognition*, 282.

Gerrig, R. J., & Rapp, D. N. (2004). Psychological processes underlying literary impact. *Poetics Today*, 25(2), 265-281.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of personality and social psychology*, 79(5), 701.

Greene, E., & Loftus, E. F. (1984). Solving the eyewitness problem. *Behavioral Sciences & the Law*, 2(4), 395–406. <https://doi.org/10.1002/bsl.2370020406>

Grice, H. P. (1975). Logic and Conversation. In P. Cole, & J. L. Morgan. (Eds.), *Syntax and Semantics*, Vol. 3, *Speech Acts* (pp. 41-58). New York: Academic Press.

Hafer, C. L., & Bègue, L. (2005). Experimental Research on Just-World Theory: Problems, Developments, and Future Challenges. *Psychological Bulletin*, 131(1), 128–167. <https://doi.org/10.1037/0033-2909.131.1.128>

Higgins, E. T. (2012). Regulatory focus theory. *Handbook of theories of social psychology*, 1, 483-504.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological*

Kitayama, S., Chua, H. F., Tompson, S., & Han, S. (2013). Neural mechanisms of dissonance: An fMRI investigation of choice justification. *NeuroImage*, 69, 206–212. doi:10.1016/j.neuroimage.2012.11.034

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. *Social psychology: Handbook of basic principles*, 655-701.

Kroeger, P. R. (2023). *Analyzing meaning: An introduction to semantics and pragmatics*. Language Science Press.

Kroneisen, M., Woehe, L., & Rausch, L. S. (2015). Expectancy effects in source memory: How moving to a bad neighborhood can change your memory. *Psychonomic Bulletin & Review*, 22(4), 1020-1037.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron*, 63(6), 889-901. <https://doi.org/10.1016/j.neuron.2009.07.030>

Lerner, M. (1980). *The belief in a just world: A Fundamental Delusion*. Springer.

Li, X. (2022). Text Classification of Conversational Implicatures Based on Lexical Features. *Applied Artificial Intelligence*, 36(1). <https://doi.org/10.1080/08839514.2022.2127598>

Lipkus, I. (1991). The construction and preliminary validation of a global belief in a just world scale and the exploratory analysis of the multidimensional belief in a just world scale. *Personality and Individual differences*, 12(11), 1171-1178.

Lipkusa, I. M., Dalbert, C., & Siegler, I. C. (1996). The importance of distinguishing the belief in a just world for self versus for others: Implications for psychological well-being. *Personality and social psychology bulletin*, 22(7), 666-677.

Loftus, E. F. (2003). Make-believe memories. *American Psychologist*, 58(11), 867–873. <https://doi.org/10.1037/0003-066X.58.11.867>

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12(4), 361-366.

Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: the creation of new memories. *Journal of experimental psychology: General*, 118(1), 100.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of verbal learning and verbal behavior*, 13(5), 585-589.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric annals*, 25(12), 720-725.

Loftus, G. R., & Loftus, E. F. (2019). Human memory: The processing of information. Psychology Press.

Loftus, G.R., & Loftus, E.F. (1976). Human Memory: The Processing of Information (1st ed.). Psychology Press. <https://doi.org/10.4324/9781315787145>

Lucas, T., Zhdanova, L., & Alexander, S. (2011). Procedural and distributive justice beliefs for self and others. *Journal of Individual Differences*.

Lutz, S., & Huitt, W. (2003). Information processing and memory: Theory and applications. *Educational Psychology Interactive*, 1-17.

Mahr, J. B., & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences*, 41<https://doi.org/10.1017/S0140525X17000012>

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87(3), 252–271. <https://doi.org/10.1037/0033-295X.87.3.252>

Mandler, J. M. (1990). Recall and its verbal expression. In R. Fivush & J. A. Hudson (Eds.), *Knowing and remembering in young children* (pp. 317–330). Cambridge University Press.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38, 299–337. <https://doi.org/10.1146/annurev.ps.38.020187.001503>

Marsh, E. J., Meade, M. L., & Roediger III, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language*, 49(4), 519-536.

Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language*, 58(2), 449-464.

Mattarozzi, K., Todorov, A. & Codispoti, (2015). M. Memory for faces: the effect of facial appearance and the context in which the face is encountered. *Psychological Research* 79, 308–317. <https://doi.org/10.1007/s00426-014-0554-8>

Mazzoni, G., & Memon, A. (2003). Imagination can create false autobiographical memories. *Psychological Science*, 14(2), 186-188.

McAdams, D. P. (2001). The psychology of life stories. *Review of general psychology*, 5(2), 100-122.

McKinney, W. (2010, June). Data structures for statistical computing in Python. In *SciPy* (Vol. 445, No. 1, pp. 51-56).

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior*, 15(4), 447-458. [https://doi.org/10.1016/S0022-5371\(76\)90040-2](https://doi.org/10.1016/S0022-5371(76)90040-2)

Murphy, G., Loftus, E. F., Grady, R. H., Levine, L. J., & Greene, C. M. (2019). False Memories for Fake News During Ireland's Abortion Referendum. *Psychological Science*, 095679761986488. doi:10.1177/0956797619864887

Murphy, S. T., Monahan, J. L., & Glaser, J. (2019). Political fake news is correlated with belief in fake news, not political ideology. *Social Science & Medicine*, 240, 112557.

Murre, J. M., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PloS one*, 10(7), e0120644.

Neuschatz, J. S., Lampinen, J. M., Preston, E. L., Hawkins, E. R., & Toglia, M. P. (2002). The effect of memory schemata on memory and the phenomenological experience of naturalistic situations. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 16(6), 687-708.

Neves, G., Cooke, S. F., & Bliss, T. V. (2008). Synaptic plasticity, memory, and the hippocampus: a neural network approach to causality. *Nature reviews neuroscience*, 9(1), 65-75.

Nieuwland, M. S., & Van Berkum, J. J. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain research*, 1118(1), 155-167.

Polišenská, K., Chiat, S., Comer, A., & McKenzie, K. (2014). Semantic effects in sentence recall: The contribution of immediate vs delayed recall in language assessment. *Journal of communication disorders*, 52, 65-77.

Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In *Advances in experimental social psychology* (Vol. 20, pp. 297-340). Academic Press.

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental review*, 31(2-3), 180-206.

Rodriguez, D. N., & Strange, D. (2015). False memories for dissonance inducing events. *Memory*, 23(2), 203–212. <https://doi.org/10.1080/09658211.2014.881501>

Roediger, H. L., & Guynn, M. J. (1996). Retrieval Processes. *Memory*, 197–236. doi:10.1016/b978-012102570-0/50009-4

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803-814.

Rubin, Z., & Peplau, L. A. (1975). Who believes in a just world?. *Journal of social issues*, 31(3), 65-89.

Schacter, D. L. (1989). On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 355–389). Lawrence Erlbaum Associates, Inc.

Schacter, D. L. (2001). *The Seven Sins of Memory: How the Mind Forgets and Remembers*.

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, 44(1), 149-160.

Schacter, D. L., & Tulving, E. (1994). *Memory Systems 1994*. MIT Press

SCHACTER, D. L., CHIAO, J. Y., & MITCHELL, J. P. (2003). The Seven Sins of Memory. *Annals of the New York Academy of Sciences*, 1001(1), 226–239. doi:10.1196/annals.1279.012

Schacter, D. L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, 35(2), 319-334.

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, 6, 379-393.

Seamon, J. G., Lee, I. A., Toner, S. K., Wheeler, R. H., Goodkind, M. S., & Birch, A. D. (2002). Thinking of critical words during study is unnecessary for false memory in the Deese, Roediger, and McDermott procedure. *Psychological Science*, 13(6), 526-531.

Sedikides, C., Green, J. D., & Pinter, B. (2004). Self-protective memory. In D. R. Beike, J. M. Lampinen, & D. A. Behrend (Eds.), *The self and memory* (pp. 161–179). Psychology Press.

Skowronski, J. J., Betz, A. L., Thompson, C. P., & Shannon, L. (1991). Social memory in everyday life: Recall of self-events and other-events. *Journal of Personality and Social Psychology*, 60(6), 831–843. <https://doi.org/10.1037/0022-3514.60.6.831>

Sperber, D. & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language* 17(1/2), 3–23

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual review of psychology*, 44, 453–495. <https://doi.org/10.1146/annurev.ps.44.020193.002321>

Stangor, C. and Walinga, J. (2014). *Introduction to Psychology* – 1st Canadian Edition. Victoria, B.C.: BCcampus. Retrieved from <https://opentextbc.ca/introductiontopsychology/>

Starmans, C., & Friedman, O. (2012). The folk conception of knowledge. *Cognition*, 124(3), 272-283.

Tajfel, H. and Turner, J.C. (1986). The Social Identity Theory of Intergroup Behavior. In: Worchel, S. and Austin, W.G., Eds., *Psychology of Intergroup Relation*, Hall Publishers, Chicago, 7–24.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tobin, S. J., & Raymundo, M. M. (2009). Persuasion by causal arguments: The motivating role of perceived causal expertise. *Social Cognition*, 27(1), 105-127.

Tonegawa, S., Liu, X., Ramirez, S., & Redondo, R. (2015). Memory engram cells have come of age. *Neuron*, 87(5), 918-931.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson, *Organization of memory*. Academic Press.

Tulving, E. (1983). Elements of episodic memory.

Tulving, E., & Craik, F. I. M. (2000). *The Oxford Handbook of Memory*. Oxford : Oxford University Press.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior*, 5(4), 381–391. [https://doi.org/10.1016/S0022-5371\(66\)80048-8](https://doi.org/10.1016/S0022-5371(66)80048-8)

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: gist-based processes in normal aging. *Psychology and aging*, 13(2), 230.

Van Rossum, G. (2007, June). Python Programming Language. In USENIX annual technical conference (Vol. 41, No. 1, pp. 1-36).

Walker, W. R., & Skowronski, J. J. (2009). The Fading affect bias: But what the hell is it for? *Applied Cognitive Psychology*, 23(8), 1122–1136. doi:10.1002/acp.1614

Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way!. *Review of general psychology*, 7(2), 203-210.

Wilson, A., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory*, 11(2), 137-149. Wright, D. B., & Loftus, E. F. (1999). Eyewitness memory: What can the law learn from psychology?. *Trends in cognitive sciences*, 3(2), 53-60.

Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind*, 111(443), 583-632. <https://doi.org/10.1093/mind/111.443.583> doi:10.1016/s0022-5371(72)80001-x

Wilson, D., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge: Cambridge University Press.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3), 441–517. <https://doi.org/10.1006/jmla.2002.2864>

Zheng, Z., Qiu, S., Fan, L., Zhu, Y., & Zhu, S. C. (2021, August). Grice: A grammar-based dataset for recovering implicature and conversational reasoning. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021* (pp. 2074-2085).

Appendix A - Vignettes

Vignette 1 - Wedding

Jane and Karen are close friends and live in different cities. Karen decided to get married and arranged a date for the wedding. She thus invited her close friends one by one. She called Jane and stated:

“I would be glad to see you at my wedding next Wednesday.”

Jane answered:

“I am going to buy a ticket for next Wednesday.”

Kept Promise	Broken Promise
On the wedding day, Jane was there.	On the wedding day, Jane was not there.

Vignette 2 - Concert

Olivia was the singer of a small punk band. They had a gig roughly every other month, and her friend Peter always joined. Peter recently became acquainted with a woman who happens to be a discographic agent. At the end of November, Olivia asked Peter if he could bring her along to her next concert:

“It would be so cool if she could see us on stage. We can speed up and have new songs ready for when she shows up.”

Peter answered:

“I’ll see her later!”

Kept Promise	Broken Promise
It turned out that Peter invited his friend, the discographic agent, to come to the concert.	It turned out that Peter did not invite his friend, the discographic agent, to come to the concert.

Vignette 3 - Cinema

John and Carol were film school students who liked to watch new movies in the cinema. John saw a new movie poster in the newspaper and asked Carol:

“Would you like to go to the cinema tomorrow?”

Carol replied:

“I will check the tickets.”

Kept Promise	Broken Promise
The evening after, it turned out that Carol bought the tickets for her and John.	The evening after, it turned out that Carol did not buy the tickets for her and John.

Vignette 4 - Plants

Bob was preparing to leave for a week-long business trip, and he was looking for someone to water the plants in his office while he was gone. Bob's plants needed to be watered quite frequently; otherwise, they would dry out. He asked his officemate Jim:

"I am off for a week; can you please water my plants while I am away?"

Jim replied:

"I will go to the office every day!"

Kept Promise	Broken Promise
When Bob came back from his trip, he discovered that Jim watered the plants.	When Bob came back from his trip, he discovered that Jim did not water the plants.

Vignette 5 - Drink

Sarah and Tom are friends, discussing their plans for the evening.

Sarah asked:

“Do you want to join me for a drink at the pub after work?.” Tom replied:

"I don't have any other plans tonight!".

Kept Promise	Broken Promise
She arrived at the bar, ordered a drink, and waited. After ten minutes, Tom arrived at the bar.	She arrived at the bar, ordered a drink, and waited. After an hour, Tom did not show up.

Vignette 6 - Draft

Werner was working on his master's thesis and was almost done with the final draft. He asked his colleague Joe to proofread the draft.

"Can you help me out and check my writing?"

Joe answered:

"I have some free time tomorrow."

Kept Promise	Broken Promise
The next day, Werner received the proofread document from Joe.	The next day, Werner did not receive the proofread document from Joe.

Vignette 7 - The Lunch

Steve and Josh were having lunch together, and both were carrying their laptops. During lunch, Steve said:

“I have to go to the restroom.”

Josh replied while pointing at Steve’s laptop:

“You can leave it here.”

Kept Promise	Broken Promise
When Steve returned from the restroom, Josh was still sitting at the table while keeping an eye on the laptop.	When Steve returned from the restroom, Josh was at the counter ordering coffee while Steve’s laptop was left on the table unattended.

Vignette 8 - The Doggie

Greg and Anna have been living together for several years. A couple of years ago, they decided to adopt a puppy. When Greg arrives home from work in the evening, he asks Anna:

“Should I go out, or did you walk the doggie already?”

Anna replies:

“I went out already.”

Kept Promise	Broken Promise
Later that day, Greg learned that Anna walked the dog half an hour before Greg arrived home.	Later that day, Greg learned that Anna did not walk the dog before Greg arrived home.

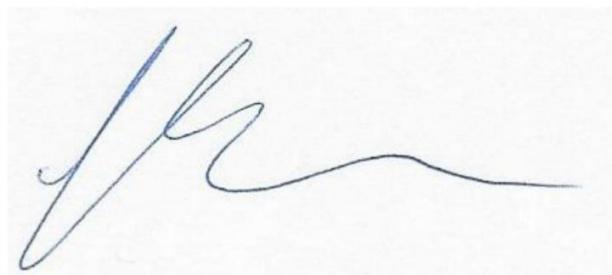
Appendix B – Ethical Approval

Psychological Research Ethics Board (PREBO)

Ethics Approval

The Psychological Research Ethics Board (PREBO) has evaluated the ethical aspects of the following research proposal, and has approved it on behalf of *Central European University*:

<i>Researcher</i> Christophe Heintz
<i>Supervisor</i>
<i>Project title</i> Motivated recall of commitments made via implicatures
<i>Reference number</i> 2023-34
<i>Date of approval</i> Sept 18, 2023

A handwritten signature in blue ink, appearing to read 'GK'.

Gunther Knoblich

Representative of PREBO

Appendix C – Attention Check Questions

Vignette 1 - Wedding

Why did Karen phone Jane?

- Karen invited Jane to join a sailing trip.
- Karen wanted to know if Jane was married.
- Karen invited Jane to her marriage, which is taking place “next Wednesday.”
- Karen wanted Jane to buy a ticket for a concert, which is taking place “next Wednesday.”

Vignette 5 - Drink

What would Sarah like to do after work?

- Stay at the office and have a drink there.
- Spend some time at the pub with Thom.
- Get back home and phone her mother.
- Get back home and phone Thom.

Vignette 6 - Draft

What does Werner want from Joe?

- Werner would like Joe to proofread the draft of his Master's thesis.
- Werner would like Joe to take some free time to travel and enjoy life.
- Werner would like Joe to write a letter to his mother.
- Werner would like Joe to read ‘Crime and Punishment.’

Vignette 7 - The Laptop

What is Steve leaving at the table when he goes to the restroom?

- Steve leaves his laptop
- Steve leaves his dog, called Josh
- Steve leaves the bill for Josh to pay
- Steve does not leave the table, it is Josh who goes to the bathroom

Vignette 8 - The Doggie

What did Greg want to know from Anna?

- whether the dog has already had his evening walk
- whether Anna wanted to adopt another dog
- whether Anna thought it was cold outside
- whether Anna wanted to split

Appendix D – Consent Form

We appreciate your interest in our study at the Central European University's Department of Cognitive Science.

Our research aims to understand how people evaluate agents' behaviours in various contexts. Participation involves reading short stories, and responding to short questions in a follow-up study.

This is a voluntary study, and we estimate it will take less than 5 minutes of your time. While we don't foresee any negative impacts, please be aware that confidentiality risks exist with any online activity. We'll do our utmost to keep your responses confidential, and no individual data will be published. We won't ask for any personal information that could identify you or put you at risk.

Your data will be stored for 10 years. For more details about your rights, please visit <https://www.ceu.edu/privacy>. For further information, feel free to contact the researcher at ustak@ceu.edu.

By ticking the box below, you're providing informed consent to participate and confirming you're over 18.

Appendix E – GitHub Repository

Main Repository

<https://github.com/KutlayUs/Imlicature-Scoring-Algorithm.git>

Experiment Data

https://github.com/KutlayUs/Imlicature-Scoring-Algorithm/blob/b76fd68536f2d40a45c48ef572da766ce8a83fab/New_CombinedOutput_DelayNoDelay.xlsx

Implicature Scoring Algorithm Source Code

https://github.com/KutlayUs/Imlicature-Scoring-Algorithm/blob/e1fa71550d184656dbf8e2884c3a5f5e7085b745/ImplicatureScoringAlgorithm_newest.py

Diagnostic Words

https://github.com/KutlayUs/Imlicature-Scoring-Algorithm/blob/c76b29b3492db67fc97ecd394af2461992e4ed29/survey_diagnostics.xls

Stopwords

https://github.com/KutlayUs/Imlicature-Scoring-Algorithm/blob/c76b29b3492db67fc97ecd394af2461992e4ed29/stop_words_english.txt

Reliability Test Data

https://github.com/KutlayUs/Imlicature-Scoring-Algorithm/blob/c76b29b3492db67fc97ecd394af2461992e4ed29/output_reliability_updated.xls

Appendix F – Belief in a Just World Scale

Lucas, T., Zhdanova, L., & Alexander, S (2011). Procedural and distributive justice beliefs for self and others: Assessment of a four-factor individual differences model. *Journal of Individual Differences*, 32, 14-25.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree

Your Rating	Statement
	1. I feel that people generally earn the rewards and punishments that they get in this world.
	2. People usually receive the outcomes that they deserve.
	3. People generally deserve the things that they are accorded.
	4. I feel that people usually receive the outcomes that they are due.
	5. People usually use fair procedures in dealing with others.
	6. I feel that people generally use methods that are fair in their evaluations of others.
	7. Regardless of the specific outcomes they receive, people are subjected to fair procedures.
	8. People are generally subjected to processes that are fair.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree

Your Rating	Statement
	1. I feel that I generally earn the rewards and punishments that I get in this world.
	2. I usually receive the outcomes that I deserve.
	3. I generally deserve the things that I am accorded.
	4. I feel that I usually receive the outcomes that I am due.
	5. People usually use fair procedures in dealing with me.
	6. I feel that people generally use methods that are fair in their evaluations of me.
	7. Regardless of the specific outcomes I receive, I am subjected to fair procedures.
	8. I am generally subjected to processes that are fair.

Appendix G – Participant Distribution

Table of Participant Distribution

Validity	Delay Condition	Commitment Condition	Vignette						Total
			Doggie	Draft	Drink	Laptop	Wedding		
Error	Delay	Broken	6	31	13	4	17	71	
		Kept	5	17	13	3	10	48	
	Delay Total		11	48	26	7	27	119	
	No Delay	Broken	2	1	4		4	11	
		Kept	1	4	3	1	2	11	
	No Delay Total		3	5	7	1	6	22	
Error Total			14	53	33	8	33	141	
Valid	Delay	Broken	30	16	16	28	16	106	
		Kept	32	15	13	27	31	118	
	Delay Total		62	31	29	55	47	224	
	No Delay	Broken	32	34	34	40	36	176	
		Kept	29	34	15	39	35	152	
	No Delay Total		61	68	49	79	71	328	
Valid Total			123	99	78	134	118	552	
Grand Total			137	152	111	142	151	693	

Appendix H – Detailed Analysis of Vignettes

For better insight, we analyzed the mean implicature and accuracy scores by vignette (Doggie, Draft, Drink, Laptop, Wedding), commitment (Broken vs. Kept), and delay (Delay vs. No Delay) conditions.

Implicature Score and Commitment Condition

The mean implicature score for the Doggie vignette was higher in the Kept condition ($M = 0.44$) than in the Broken condition ($M = 0.25$). Similarly, the Draft vignette showed higher implicature scores in the Kept condition ($M = 0.31$) than in the Broken condition ($M = 0.21$). The Drink vignette followed this pattern with higher implicature scores in the Kept ($M = 0.33$) than in the Broken condition ($M = 0.25$). The Laptop vignette had the highest implicature score in the Kept condition ($M = 0.53$) compared to the Broken ($M = 0.38$). However, for the Wedding vignette, the implicature score was slightly higher in the Broken condition ($M = 0.41$) than in the Kept condition ($M = 0.32$).

Accuracy Score and Commitment Condition

In the Doggie vignette, the Broken condition ($M = 0.73$) resulted in a higher mean accuracy score than the Kept condition ($M = 0.61$). For the Draft vignette, the Kept condition ($M = 0.68$) had a marginally higher accuracy score than the Broken condition ($M = 0.65$). The Drink vignette showed a higher mean accuracy in the Broken condition ($M = 0.58$) versus the Kept condition ($M = 0.53$). In contrast, the Laptop vignette had a higher accuracy score in the Kept condition ($M = 0.73$) than in the Broken condition ($M = 0.59$). The Wedding vignette showed slightly higher accuracy in the Broken condition ($M = 0.73$) than in the Kept condition ($M = 0.68$).

Implicature Score and Delay Condition

In the Doggie vignette, the No Delay condition ($M = 0.42$) showed a higher mean implicature score than the Delay condition ($M = 0.27$). For the Draft vignette, the Delay condition ($M = 0.48$) resulted in a significantly higher score than the No Delay condition ($M = 0.04$). The Drink vignette showed a higher implicature score in the No Delay condition ($M = 0.50$) than in the

Delay condition ($M = 0.15$). The Laptop vignette had the highest score in the Delay condition ($M = 0.77$) compared to the No Delay condition ($M = 0.15$). For the Wedding vignette, the No Delay condition ($M = 0.20$) showed a slightly higher score than the Delay condition ($M = 0.15$).

Accuracy Score and Delay Condition

In the Doggie vignette, the No Delay condition ($M = 0.87$) had a higher mean accuracy score than the Delay condition ($M = 0.58$). The Draft vignette showed a similar pattern with the No Delay condition ($M = 0.87$), surpassing the Delay condition ($M = 0.46$). For the Drink vignette, the No Delay condition ($M = 0.77$) resulted in a higher mean accuracy score than the Delay condition ($M = 0.34$). The Laptop vignette had a higher accuracy score in the No Delay condition ($M = 0.67$) than in the Delay condition ($M = 0.46$). The Wedding vignette showed higher accuracy in the No Delay condition ($M = 0.70$) compared to the Delay condition ($M = 0.42$).

In summary, higher mean implicature scores were generally observed in the Kept condition compared to the Broken condition, except for the Wedding vignette. There were mixed results under the delay condition, with the No Delay condition generally showing higher scores except for the Draft vignette. For mean accuracy scores, higher accuracy was mainly observed in the Broken condition for some vignettes, such as Doggie and Laptop. The No Delay condition consistently resulted in higher accuracy scores across all vignettes compared to the Delay condition. These results suggest that commitment and delay conditions significantly impact both implicature and accuracy scores across various vignettes, demonstrating different patterns depending on the vignette and condition.

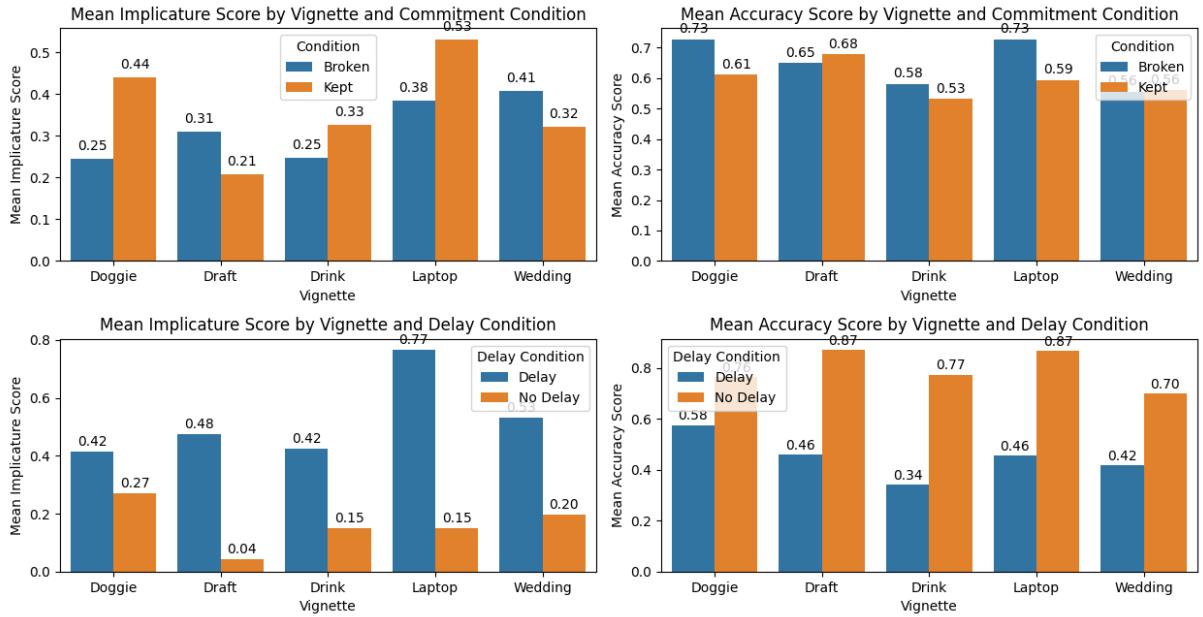


Figure 16 Comparison of Means of Implicature and Accuracy Scores

In addition, we analyzed the distribution of error-labeled participant recalls across different vignettes under varying conditions. For the commitment condition, the Doggie vignette showed a higher percentage of error-labeled recalls in the Broken condition (57%) compared to the Kept condition (43%). Similarly, the Draft vignette had more errors in the Broken condition (60%) than in the Kept condition (40%). The Drink vignette followed this trend, with 52% errors in the Broken condition and 48% in the Kept condition. The Laptop vignette presented an equal percentage of error-labeled recalls (50%) in both conditions. However, the Wedding vignette exhibited a significantly higher error rate in the Broken condition (64%) than in the Kept condition (36%).

When examining the delay condition, the Doggie vignette revealed a substantially higher percentage of error-labeled recalls in the Delay condition (79%) than in the No Delay condition (21%). The Draft vignette showed an even more pronounced difference, with 91% errors in the Delay condition and only 9% in the No Delay condition. The Drink vignette also had a higher error rate in the Delay condition (79%) versus the No Delay condition (21%). The Laptop vignette displayed a similar pattern, with 88% errors in the Delay condition compared to 13% in the No Delay condition. Lastly, the Wedding vignette showed higher error-labeled recalls in the Delay condition (82%) than in the No Delay condition (18%).

These results indicate that commitment and delay conditions significantly impact the error rates in participant recalls. The Broken condition consistently resulted in higher error rates across all vignettes than the Kept condition, with the most notable difference observed in the Wedding vignette. Similarly, the Delay condition consistently led to higher error rates across all vignettes than the No Delay condition, with the Draft vignette showing the most notable difference.

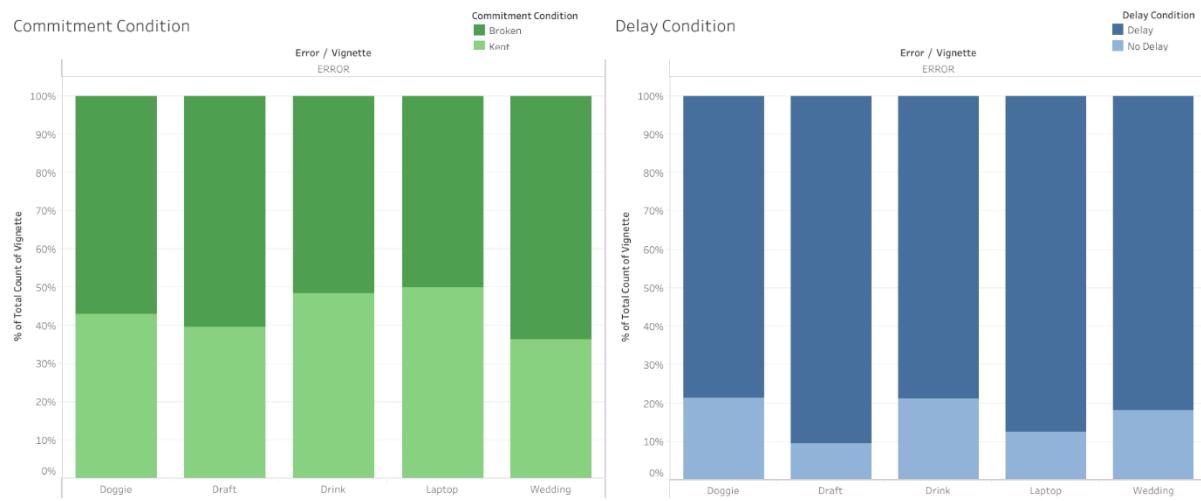


Figure 17 The Percentage of the Total Count of Error for Each Vignette

Discussion

Our study analyzed the mean implicature and accuracy scores across different vignettes under varying commitment conditions (Kept vs. Broken) and delay (No Delay vs. Delay). Each vignette presented distinct scenarios, leading to varied results based on their emotional salience, narrative complexity, and contextual relevance.

Implicature Score and Commitment Condition

The Doggie vignette, involving the emotional context of pet care, showed higher implicature scores in the Kept condition ($M = 0.44$) compared to the Broken condition ($M = 0.25$). This suggests that fulfilling a promise related to pet care reinforces participants' ability to make inferences, likely due to the task's strong emotional engagement and routine nature. The predictability and emotional attachment in this context enhances cognitive processing when promises are kept.

The Draft vignette involving academic commitments showed better implicature scores when the commitment was kept ($M = 0.31$ vs. $M = 0.21$). This indicates that a fulfilled professional promise enhances cognitive processing and understanding, possibly due to the reduced cognitive load and clear narrative flow. The structured and professional context likely aids in better comprehension and recall.

As illustrated in the Drink vignette, social promises resulted in higher implicature scores when kept ($M = 0.33$) than broken ($M = 0.25$). Social engagements likely carry emotional weight, and keeping such promises helps better understand and recall. The interpersonal nature of this scenario makes it easier to process and remember when the promise is fulfilled.

The Laptop vignette showed the highest implicature score in the Kept condition ($M = 0.53$) compared to the broken condition ($M = 0.38$), suggesting that promise fulfillment benefits professional and technical contexts by enhancing clarity and cognitive processing. The clear and direct consequences of the actions involved make it easier for participants to process the commitment when it is kept.

Interestingly, the Wedding vignette had higher implicature scores in the Broken condition ($M = 0.41$) than in the Kept condition ($M = 0.32$). The emotional salience of a wedding and the significant impact of a broken promise in this context might heighten emotional and cognitive responses, leading to better inferential understanding despite the negative outcome. The high stakes and strong emotional context of a wedding may cause broken promises to stand out more in memory.

Accuracy Score and Commitment Condition

Regarding accuracy scores, the Doggie vignette showed higher accuracy in the Broken condition ($M = 0.73$) compared to the Kept condition ($M = 0.61$), suggesting that the emotional dissonance of unfulfilled pet care responsibilities enhances memory retention. The emotional context increases the cognitive load and retention when promises are broken.

There was a minimal difference in accuracy scores for the Draft vignette, with the Kept condition slightly higher ($M = 0.68$) than the Broken condition ($M = 0.65$). This indicates that professional commitments are well remembered regardless of the outcome. Academic commitments' importance and structured nature ensure good recall in both conditions.

The Drink vignette also showed higher accuracy in the Broken condition ($M = 0.58$) compared to the Kept condition ($M = 0.53$), possibly due to the emotional impact of an unmet social promise. The interpersonal disappointment likely enhances memory retention.

The Laptop vignette showed higher accuracy when the promise was kept ($M = 0.73$) than when it was broken ($M = 0.59$), indicating that promise fulfillment benefits professional and security-related commitments. The clear and direct outcomes in a professional setting likely aid in better recall when promises are kept.

In the Wedding vignette, higher accuracy was observed in the Broken condition ($M = 0.73$) compared to the Kept condition ($M = 0.68$), highlighting the strong emotional impact of a broken wedding promise. The high emotional stakes of a wedding ensure better recall of broken promises.

Implicature Score and Delay Condition

The Doggie vignette showed higher implicature scores in the No Delay condition ($M = 0.42$) than in the Delay condition ($M = 0.27$), suggesting that immediate recall improves implicature understanding due to fresh emotional engagement. The routine and emotional nature of pet care likely benefit from immediate processing.

Conversely, the Draft vignette showed higher implicature scores with a delay ($M = 0.48$) than no delay ($M = 0.04$), suggesting that participants need more time to process and understand professional commitments. The complexity and professional nature may require more cognitive processing time.

The Drink vignette showed higher implicature scores in the No Delay condition ($M = 0.50$) compared to the Delay condition ($M = 0.15$), indicating that social promises benefit from immediate recall. The social and immediate nature of the context likely benefits from quick processing.

The Laptop vignette showed significantly higher implicature scores with a delay ($M = 0.77$) than no delay ($M = 0.15$), suggesting that technical and professional scenarios benefit from additional processing time. The context's detailed and technical nature may require more cognitive processing time.

The Wedding vignette showed minimal difference between conditions, with slightly higher implicature scores in the No Delay condition ($M = 0.20$) compared to the Delay condition ($M = 0.15$). The high emotional context of a wedding may ensure good recall regardless of delay.

Accuracy Score and Delay Condition

For accuracy scores, the Doggie vignette showed higher accuracy in the No Delay condition ($M = 0.87$) compared to the Delay condition ($M = 0.58$), emphasizing the importance of immediate reinforcement for emotional and routine tasks. Immediate recall benefits emotional and routine tasks like pet care.

Similarly, the Draft vignette showed higher accuracy with no delay ($M = 0.87$) than delay ($M = 0.46$), indicating that professional commitments show better recall without delay. The structured and professional context benefits from immediate recall.

The Drink vignette also showed higher accuracy in the No Delay condition ($M = 0.77$) than in the Delay condition ($M = 0.34$), suggesting that social engagements benefit from immediate recall and quick processing of the social and emotional context.

The Laptop vignette showed higher accuracy in the No Delay condition ($M = 0.67$) than in the Delay condition ($M = 0.46$), highlighting the importance of prompt recall for professional and security-related scenarios. Immediate recall benefits the professional and detailed context.

The Wedding vignette also showed higher accuracy in the No Delay condition ($M = 0.70$) than the Delay condition ($M = 0.42$), reinforcing the significance of immediate recall in highly emotional contexts. The high stakes and emotional nature of a wedding benefit from immediate processing.

Error-Labeled Participant Recalls

The analysis of error-labeled participant recalls across different conditions revealed significant impacts of both commitment and delay conditions. For the commitment condition, the Doggie vignette showed a higher percentage of error-labeled recalls in the Broken condition (57%) compared to the Kept condition (43%). Similarly, the Draft vignette had more errors in the Broken condition (60%) than in the Kept condition (40%). The Drink vignette followed this trend, with 52% errors in the Broken condition and 48% in the Kept condition. The Laptop

vignette presented an equal percentage of error-labeled recalls (50%) in both conditions. However, the Wedding vignette exhibited a significantly higher error rate in the Broken condition (64%) than in the Kept condition (36%).

When examining the delay condition, the Doggie vignette revealed a substantially higher percentage of error-labeled recalls in the Delay condition (79%) than in the No Delay condition (21%). The Draft vignette showed an even more pronounced difference, with 91% errors in the Delay condition and only 9% in the No Delay condition. The Drink vignette also had a higher error rate in the Delay condition (79%) versus the No Delay condition (21%). The Laptop vignette displayed a similar pattern, with 88% errors in the Delay condition compared to 13% in the No Delay condition. Lastly, the Wedding vignette showed higher error-labeled recalls in the Delay condition (82%) than in the No Delay condition (18%).

These results indicate that commitment and delay conditions significantly impact the error rates in participant recalls. The Broken condition consistently resulted in higher error rates across all vignettes than the Kept condition, with the most notable difference observed in the Wedding vignette. Similarly, the Delay condition consistently led to higher error rates across all vignettes than the No Delay condition, with the Draft vignette showing the most pronounced difference. These findings suggest that errors in participant recalls are more prevalent under broken commitments and delay conditions.

The results indicate that the nature of each vignette, including emotional salience, narrative complexity, and contextual relevance, plays a crucial role in how participants process and recall information. Emotional and routine tasks like the Doggie and Drink vignettes benefit from immediate recall and emotional engagement. In contrast, professional commitments, such as the Draft and Laptop vignettes, may benefit from additional processing time. The Wedding vignette's unique emotional impact highlights the complexity of memory processes, where emotional salience can significantly enhance recall even in the face of broken promises.

Future research should consider these factors when designing studies and interpreting results, providing valuable insights into the mechanisms underlying memory formation and recall in different contexts. Understanding these nuances can help tailor interventions and strategies to improve memory retention based on different scenarios' emotional and cognitive demands.