PREFERENCES, TIMING, AND LIMITATIONS DURING REASONING: A DUAL PROCESS APPROACH

BURAK BERK GÜLTEKİN

JUNE, 2022

PREFERENCES, TIMING, AND LIMITATIONS DURING REASONING: A DUAL PROCESS APPROACH

BY

BURAK BERK GÜLTEKİN

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN

COGNITIVE SCIENCES

YEDİTEPE UNIVERSITY

JUNE, 2022

PLAGIARISM

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Date: 16.06.2022

Name/Surname: Burak Berk Gültekin

Signature:

ÖZET

Akıl yürütme süreçleri, sezgisel sistem ve analitik sistem tarafından nasıl şekillenir? Bu çalışmada, sezgilerin ve analitik düşünmenin akıl yürütme sürecinde verilen yanıt üzerinde nasıl etkileri olduğuna odaklandık. Çift süreç teorilerine göre akıl yürütme, Sistem 1 ve Sistem 2 süreçleri olmak üzere iki farklı işlem modundan oluşur. Sistem 1 otomatik, hızlı ve içgüdüsel yanıtların üretilmesinde etkili olan sistemdir. Sistem 1 çalışma belleği kaynaklarına ihtiyaç duymaz. Diğer sistem mantığa dayanan sistemdir. Bu sistemin özellikleri temkinli, yavaş ve bilinçlidir. Sistem 2 çalışma belleği kapasitesine bağlıdır. Bu nedenle, sistem 2'nin çalışma belleği kaynaklarına ihtiyacı vardır. Bu çalışmada, deney tasarımı şu şekildedir: Tüm katılımcılar, ilk yanıt aşamasında bellek yükünü akıllarında tutarken ve bir zaman sınırına sahipken, rasgele dağıtılmış sopa ve top problemlerinin değiştirilmiş versiyonunu ve kıyassal akıl yürütme problemlerini yanıtladılar. Daha sonra sorular bellek yükü ve süre sınırı olmadan tekrar soruldu. Her iki görevde çelişki içeren (deney soruları) ve çelişki içermeyen (kontrol soruları) sorulardan oluşur. Katılımcılar sopa ve top problemlerinin değiştirilmiş bir versiyonunda dört seçenek arasından seçim yaparken, kıyassal akıl yürütme problemlerinde iki öncül ve bir sonucu göz önünde bulundurarak evet ve hayır arasında seçim yapmışlardır. İki yanıt paradigması uygulandı. Problemler iki defa soruldu. Problemlerin ilk soruluşunda, ikili görev prosedürü ve zaman baskısı ile çalışma belleği meşgul edildi. Sistem 1 ile hızlı düşünme ve sistem 2 ile yavaş düşünme yaygın çıkarımını esas aldık.

Anahtar Kelimeler: akıl yürütme, analitik sistem, çift süreç teorileri, sezgisel sistem, çelişki algılaması, çalışma belleği, iki yanıt paradigması, sezgiler & yanlılıklar

ABSTRACT

How are reasoning processes shaped by the heuristic system and analytic system? In the present study, we were focused on how heuristics and analytical thinking affect reasoning. According to dual-process theories, reasoning consists of two distinct modes of processing: System 1 and System 2. System 1 generates responses that are automatic, fast, and intuitive. This system does not need working memory resources. The other system is reflective. Characteristics of this system are deliberative, slow, and conscious. System 2 depends on working memory capacity. Therefore, system 2 needs working memory resources. In the present study, the experimental design was as follows: All participants performed randomly distributed modified version of the bat and ball problems and the syllogistic reasoning problems while they kept memory load in mind and had a time limit in the initial response stage. Then the questions were asked again without memory load and deadline. Both tasks consist of conflict (experiment questions) and no-conflict (control questions) questions. Whereas participants chose among four options in a modified version of the bat and ball problems, they chose between yes and no, considering two premises and a conclusion in the syllogistic reasoning problems. A two-response paradigm was applied. The problems were asked twice. In the first presentation of the problems, we occupied the participant's working memory with a dual task procedure and time pressure. We replicated the common implication of system 1 with fast response and system 2 with slow thinking.

Keywords: analytic system, conflict detection, dual process of thinking, heuristics & biases, reasoning, two response paradigm, working memory

To My Mother…

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to my supervisor Assistant Prof. Dr. Sara Saban for her guidance, encouragement, and feedback throughout the research. She guided me with her valuable comments. I appreciate her contributions.

I would like to thank Assistant Prof. Dr.Serkan Şener. I have learned a lot about psycholinguistics from him. Ajda Yılmaz made invaluable comments and allowed me to broaden my perspective. I express my deepest gratitude to Ajda Yılmaz. I also thank my friends Emre Gözütok, Esra Cesur and Atakan Kesim, who motivated me. I thank my family for their sincere support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1. Background

The brain can be thought as a device whose abilities including motor activities, perception, and memory have evolved to reinforce the decision that regulates actions (Gazzaniga et al., 2008). The first question is why and how do people make decisions? When a person comes across choices, he or she selects an available course of action (Smith, E. E., & Kosslyn, S. M., 2007).

In daily events, many decisions are easy to make even it contains uncertainty since some factors might be very important for the decision maker who can choose fast in decision making. When one option is dominated by the other option, even if lots of possibilities he faces, the decision maker selects the option which is highly valuable for him (Hastie & Dawes, 2010). Indeed, many decisions are not taken that easily. When values and outcomes are undefined, it becomes more difficult to make a decision. Three elements, which constitute a decision, are alternatives, consequences, and beliefs (Smith, E. E., & Kosslyn, S. M., 2007). Alternatives include several strategies, options, and choices. These are available courses of action that reasoners can consider. The benefits and losses, which are consequences, follow the choice of a particular alternative. There are three determiners, which are “outcomes, values, and utilities”, which specify consequences. Whereas the outcome belongs to the result, the value is equal to its net worth, and the utilities are considered as the desirability of the value (Smith, E. E., & Kosslyn, S. M., 2007).

Normative prescriptive theories focus on how people should reason. Consequences and alternatives are evaluated in terms of probability rules and logical principles (Baron, 2012).

People’s natural judgments do not always depend on the rules and principles. Therefore, psychologically descriptive models investigate how people make decisions (Evans, 2003).

Decision-making is a subject, which has been studied in several areas such as behavioral economics, and cognitive science. Decision-making is the mental process from the perspective of cognitive psychology (Smith, E. E., & Kosslyn, S. M., 2007).

Kahneman and Tversky examined the role of intuitions in reasoning and decision-making. They found fallacies that were made by the reasoners. People’s intuitions are shaped by past experiences and mental models including categories. The reason why categories and stereotypes, which are produced by culture, play a role in predictions is that they are formed over similarities and patterns (Tversky & Kahneman, 1983).

Tversky and Kahneman gave an example from a representative sample: "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.". After giving the information about Steve, the question was asked as “Is Steve more likely to be a librarian or a farmer?” (Tversky & Kahneman, 1974). Steve’s personality showed similarity to the stereotype of a librarian (Kahneman, 2011). This thought relies on occupational stereotypes. However, the rates indicated that there are fewer librarians than farmers in the population. When the options were evaluated by representativeness, the base rate was neglected.

De Neys (2006a) presented four experiments with four different purposes. Experiment 1 was designed for a chronometric comparison between heuristic response and analytic response. Four tasks, which were presented respectively, are “Linda Problem, Bill Problem, Drinking Age Problem, Destination Problem”. In experiment 2, the problems were presented with the executive finger tapping task that was used as an attention-demanding secondary task. Experiment 3 was planned Linda and job problems with spatial storage task that was used as working memory load. In experiment 4, the drinking age problem and card selection task were presented with a spatial storage task. The reason why De Neys (2006a) used secondary tasks to evaluate dual process predictions is that cognitive load affects system 1 and system 2 in different ways.

Linda and Bill problems are considered as conjunction fallacy. The drinking age problem is deontic reasoning. The destination problem is indicative reasoning. Although different problems were used, the conjunction fallacy and the selection task problems formed the basis of the problems in this experiment. The conjunction fallacy questions consist of short personality descriptions. The questions were formed and explained by Tversky & Kahneman (1983). Predictors read the personality descriptions of Linda and Bill. They were asked to rank statements in order of most probable to least probable. In the problem, the presented information about Linda, who is the fictitious individual, is her age, marriage status, and personal features such as “outspoken, and very bright”, major (philosophy) that she graduated from, social issues (discrimination and social justice) that she was interested in as a student. Then statements were given to rank. These are “(A) Linda is a bank teller”, and “(B) Linda is a bank teller and is active in the feminist movement” (Tversky & Kahneman, 1983). When the reader learns that Linda is interested in social issues including discrimination and social justice, Linda is representative of being active in the feminist movement. Representativeness plays a role in generating the automatic response when there is a correspondence between a sample and a category (Tversky & Kahneman, 1974). Representativeness heuristics depend on stereotypes. Kahneman and Tversky (1983) observed that most of the students (over %80) deviated from the conjunction rule. Participants ranked that statement B is more probable than statement A. Considering the conjunction rule of probability, P (A) is higher than P (B). The conjunction rule(“the probability of a conjunction of two events cannot exceed that of its constituents") is the most basic rule of probability (Tversky & Kahneman, 1983).

The other task is Wason's selection task, which is the most influential task in the psychology literature on reasoning. In this task, four cards were used. There were numbers on the one side and there were letters on the other side. Whereas participants saw the number side of the two cards, they saw the letter side of the other two cards. Therefore, A, T, 7, and 4 were seen on their visible side by participants. Participants were asked to turn over the cards to prove the rule which is "If there is an A on one side, then there is a 4 on the other side". To solve this question correctly, the "logical falsification principle" should be used. According to this principle, it is necessary to turn over the cards that falsify the given statement (If there is an A on one side, then there is a 4 on the other side). The solution ought to be the form of "If P then not Q". When the statement is not falsifiable, its truth is shown. Thus, cards A and 7 must be turned over. It was the correct answer to this question. The vast majority of participants, over %90 deviated from the logical falsification rule and gave incorrect answers (Evans et al., 2019; Manktelow & Galbraith, 2012). Most of the subjects chose cards A and 4 which were stated in the description. These cards are in the form of "If P then Q". It was called the matching bias. The matching bias in the selection task and conjunction fallacy in the Linda problem results from the heuristic system (system 1). The heuristic system produces useful solutions in many situations but it sometimes causes biased reasoning which leads to typical failure (K. E. and W. R. F. Stanovich, 2006). When briefly, in both the Linda task and the card selection task, the majority of the participants violated the logical rules and produced incorrect responses. It was shown that there is a difference between participants' performance and normative standards. The results from the study of De Neys (2006a) follow:

- In experiment 1, conjunction fallacy and indicative selection tasks were asked to the participants. The idea behind the experiment was to observe a chronometric comparison of deliberative and heuristic responses. The fast answer was labeled as a heuristic response. Because the heuristic system needs less inference time than the analytic system. The results of experiment 1 support the time course assumption of dual-process theories. The correct responses require more processing time than intuitive incorrect responses.
- In experiment 2, conjunction fallacy and indicative selection tasks with the secondary finger tapping task were presented simultaneously. A dual task procedure was applied. The participants tapped on "B, M, N, V" while they were solving the problems. It is an attention-demanding secondary task. Therefore, the central executive component of the working memory was burdened. The results show that incorrect answers increased and correct responses decreased in the conjunction fallacy questions. The secondary task affected the accuracy of responses. It was considered that the heuristic responses were given more frequently. In the card selection task, the incorrect answer, which is the matching card selections, increased. The correct logical answer, which is P and not-Q selection, decreased. Both increases and decreases are significant. There is no significant difference among the other card selections.
- In the following experiments, the reasoning tasks were applied with the secondary working memory task. The dot memory task, which is a spatial storage task (Miyake et al., 2001), was used to load the working memory while participants were answering the questions. Participants attempted to remember

presented dot patterns including three dots and four dots. The 3x3 matrix was used to fill. The results demonstrated that correct responses decreased under cognitive load in experiment 3. Dot memorization caused analytic thinking to be repressed by heuristics. In experiment 4, the incorrect matching card selections increased under the executive burden. The correct answers decreased. Unlike the experiment 2 result, the secondary memory load does not affect the deontic card selection task.

All the results are consistent with the assumptions of dual process theories in the study of De Neys (2006a).

1.2. Two Minds, Dual Process of Reasoning

Dual process theories were placed after the revolution in cognitive psychology. There are different aspects before the cognitive revolution such as the explanation of memory that is separated into implicit and explicit. Intuition and reasoning had thought subdivisions of cognitive processes from very old times. When it comes to modern times, the context of the thinking processes was described, and then dual-process theories were constituted (Kahneman & Frederick, 2012). Modern theories have described dual processing as “type 1” processing and “type 2” processing. Whereas type 1 belongs to automatic, tacit, and fast, type 2 is associated with deliberate, conscious, and slow. With the development of the dual processing approaches, the intensity of research has shifted from learning experiments to memory experiments (Frankish & Evans, 2012). Atkinson and Shiffrin (1968) categorized memory into “short term storage” and “long term storage” (Atkinson & Shiffrin, 1968). So far, type 1 and type 2 processes have been considered by distinguishing between automatic and conscious processes, respectively.

Dual system theories were developed to differentiate two discrete cognitive systems. Dual system theory explains that the cognitive reasoning system consists of two different components. One of the components, which is called system 1, is characterized as implicit, intuitive, automatic, and associative. System 1 allows people to produce a quick response. System 1 response is heavily contextualized. This response is based on prior semantic knowledge including similarities and prototypes created in the past, it comes out as a judgment (Sloman, 1996). The other component is termed System 2. The type of reasoning system is characterized as controllable, deliberative, and cautious. System 2 makes high demands on working memory. System 2 thinking process has capable of abstraction and generalization. The responses produced by system 2 are based on explicit knowledge (Kahneman, 2011).

In the psychology of reasoning and decision making, three paradigms became dominant (Evans, 2008). One of these paradigms is the “biases and heuristics” that include the judgment of probability (Gilovich et al., 2002; Kahneman & Tversky, 2013). The other research program focused on decision making under risk (Koehler & Harvey, 2008). The last paradigm is related to social judgment theory. Kahneman and Frederick (2005, 2012) focused on probability judgment. A dual process probability judgment was improved. It was found that heuristic responses lead to biases. When incorrect responses were studied, It was figured out that representativeness (Kahneman & Tversky, 1972) and availability (Tversky & Kahneman, 1973) as heuristics generate biases in probability judgment. Analytic reasoning, which might prevent these biases, is associated with System 2. Therefore, it seems that normatively correct responses differ from biased responses. Similar differentiation was observed in deductive reasoning (Evans, 2008).

1.3. Serial Default Interventionist Dual Process Theory and Alternative Interpretations of The Dual-System and Dual-Process Models

Jonathan Evans (2006), Daniel Kahneman (2011), and Keith Stanovich (2006) contributed to the massive development of the Default-interventionist (DI) model. DI model explains the role of two different systems through mental activity. DI model describes not only the properties of the two systems but also the interaction between system 1 and system 2. Serial DI model supposes that the reasoning process begins with system 1 which is thought of as a default system (De Neys, 2017). This system generates an answer fast. When the reasoner prefers an option fast with minimum cognitive effort, system 1 is engaged to choose an option. The preference of a cognitive miser depends on system 1 (De Neys, 2017). Because a decision, which is produced by system 2, demands computational resources. Therefore, system 2 activation needs more cognitive effort than system 1 activation. Also, there is a temporal difference between activation of system 1 and activation of system 2. The first heuristic response that comes to mind is generated by system 1. In order for system 2 to be activated, the necessary cognitive effort should be made. This takes more time than the first answer that comes to mind (Kahneman & Frederick, 2012).

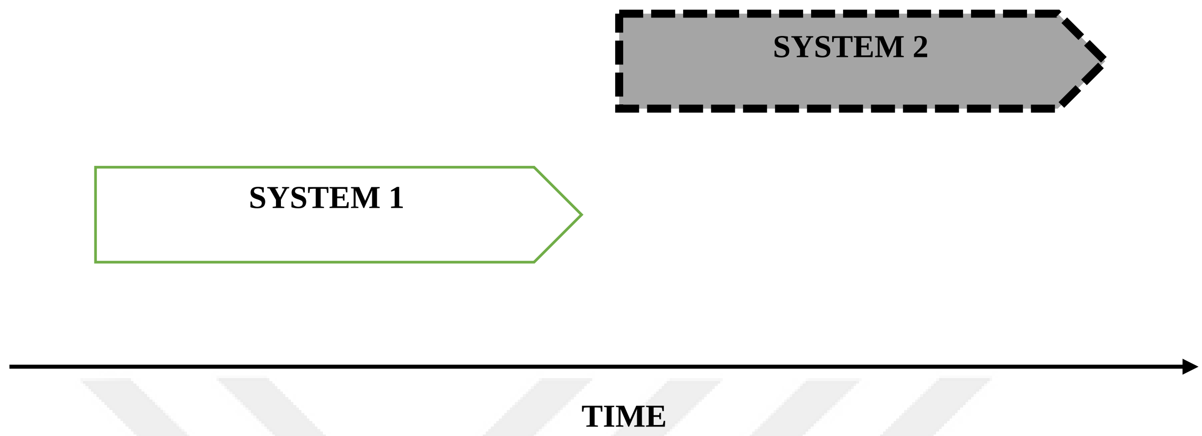

Figure 1. Illustration of Serial, Default model including the relation between System 1 and System 2 processing.

DI model assumes that system 1 is included in the reasoning processes first. Then, system 2 might be engaged. Activation of system 2 is optional. After system 1 is activated, if a necessary cognitive effort is given, system 2 engagement occurs (Evans, 2003).

Empirical findings supported the distinction between reasoning processes. Therefore, dual process theories signify two distinct cognitive processing underlying reasoning. The neutral terms, which are system 1 and system 2, were used in the literature (Evans, 2008; Kahneman & Frederick, 2012; K.E. Stanovich, 1999). Different researchers have offered various terms to name two types of thinking such as “input modules versus higher cognition” (Fodor, 2018), “adaptive unconscious versus conscious” (Wilson, 2022), and “impulsive versus reflective” (Strack & Deutsch, 2004), “stimulus bound versus higher order” (Toates, 2006), “system 1 versus system 2” (Evans, 2003; Kahneman & Frederick, 2012). Even though two distinct processing were labeled by the authors, there is not just a dual processing theory to explain thinking processing during reasoning. Some of the researchers in this field disagreed with the assumptions about "underlying cognitive systems; some propose parallel and some sequential relationships between the two processes, and so on"(Evans, 2008).

The dualistic paradigm has gained importance in judgment and decision making domains (Kahneman, 2002; Sloman, 1996; Strack & Deutsch, 2004). In the last three decades, a significant number of models have been argued based on the assumption that judgments can be formed through two qualitatively different processes or systems (Kruglanski & Gigerenzer, 2011). Some authors criticized the two systems approach. Keren and Schul (2009) proposed that the two systems are "continuous rather than dichotomous". The criticism continues as follows: There is no sequential relationship between system 1 and system 2 which are inseparable. Hence, there are no two distinct isolated processes (Keren & Schul, 2009). Reasoning and decision making researchers have termed differently to classify the two systems. Keren and Schul (2009) indicated that using several different terminologies created confusion that leads to conceptual vagueness and a lack of clarification with regard to the two-system models. Keren and Schul (2009) urged researchers to keep on a "uni-model".

Kruglanski and Gigerenzer (2011) represented a "unified theory of judgment" which might be an alternative to the dualistic paradigm. According to this theory, intuitive and deliberative judgments depend on rules. These rules determine what kind of decision is it, optimizing or satisficing. It is important to note that judgments as intuitive and deliberative do not have to arise from different rules. A rule can be the basis of intuitive and deliberative judgments. Reasoners chose the rule from their "adaptive toolbox" for a given task (Kruglanski & Gigerenzer, 2011). Kruglanski and Gigerenzer (2011) stated the factors that determine the selection of a rule as follows: The applicable rules may be limited by individual memory and the task given. The choice of a rule is guided by the perceived components as ecological rationality and the skills that an individual is able to use during the task. If different rules have approximately equal ecological rationality, it creates a conflict about which rule to choose. In this case, the chosen rule might have interfered with other competing rules. “Rules are based on core cognitive capacities, such as recognition memory” (Kruglanski & Gigerenzer, 2011). Individual differences due to varying cognitive capacities affect the application of a rule in terms of the speed and the accuracy. When intuitive judgments are evaluated with respect to rules, these judgments can be an easy or difficult rule. It is determined by the “degree of their routinization and their momentary accessibility” (Kruglanski & Gigerenzer, 2011). Therefore, this rule also might be deliberated. There is a relationship between individuals’ processing potential and ecological rationality. If the way of applying rules have repeated, it will increase the processing potential of that individual and facilitate the choice of a rule. Whether a rule is easy or difficult is related to individuals’ processing potential. If processing potential is restricted, easy-to-apply rules guide judgments (Kruglanski & Gigerenzer, 2011). More processing potential is required to apply difficult rules. Individuals, who have high processing potential, evaluate easy and difficult rules in accordance with their ecological rationality. After the evaluation, these individuals are able to apply the rule which was selected within the easy and difficult rules.

1.4. Conflict Detection

Cognitive Reflection Test (CRT) was investigated by Frederick (2005). The bat and ball problem, which is used in the CRT, became popular in the field of reasoning literature. This question was used to detect whether a participant’s response was biased or genuine. The well-known bat and problem (Frederick, 2005) follows:

“A bat and a ball together cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?” (Frederick, 2005).

The first and fast response to this question is “10 cents” (Kahneman, 2011). When a person gives an intuitive reaction, the answer becomes “10 cents”. If the response was 10 cents, the total cost would become $1.20 since the bat costs $1 more than the ball. When the cost of the ball is 10 cents, 10 cents plus $1.10, which is the cost of the bat, is equal to $1.20. However, a bat and a ball together should cost $1.10. Therefore, the intuitive “10 cents” is an incorrect heuristic answer. When a person thinks deliberatively, the correct logical answer reveals that the ball costs “5 cents”. The logic behind the correct answer is that if we convert the statement in the bat and ball problem into the equation, we can represent it as:

$$x + x + 1 = 1.10$$

$$2x = 0.10$$

$$x = 0.05$$

In the equation, x signifies the cost of the ball and x+1 demonstrates the cost of the bat. When the question is expressed in the form of an equation, the correct logical answer reveals without a doubt. The ball costs 5 cents and the bat costs $1.05. The biased reasoner’s response depends on System 1 in the problem. Therefore, intuitive responses were given (De Neys, 2017). Both conflict and no-conflict control problems of the bat and ball problem were presented in the studies. No-conflict, control problem follows:

“A bat and a ball together cost $1.10. The bat costs $1. How much does the ball cost?” (Frederick, 2005).

In this problem, there is no conflict item. Therefore, the first, and quick response produced by system 1 is correct, which is the intuitive “10 cents” response.

The comparison between reasoners' processing of the control and conflict versions can be used simply to test reasoners' conflict detection sensitivity. Response latencies and response confidence results inform the nature of the conflict detection sensitivity. Biased reasoners have longer response latencies. Also, their response confidence decreased. Biased reasoners tend to give heuristic cued system 1 response in the problems. Whereas the system 1 response is to be incorrect in the conflict problems, this response is correct in the no-conflict problems. When the reasoners have sensitivity to their incorrect response in conflict problems, they get the chance to think deliberatively about the problem and they are able to change their first intuitive response that comes to mind. Being sensitive about conflicts affects reasoning processing (De Neys & Glumicic, 2008). There are various models including the parallel model, the serial DI model, and the hybrid model to evaluate reasoning processing. Conflict detection sensitivity is the key subject for these models. The question is which system allows conflict sensitivity to develop. The researchers have dealt with this issue deactivated system 2 by applying simultaneous cognitive load and time pressure. Because it is the fact that system 2 processing is time consuming and demands cognitive sources. Bago & De Neys (2017) applied time pressure manipulation and they used concurrent load manipulation to occupy participants' cognitive sources. Therefore, system 2 becomes dysfunctional. However, some empirical (Bago & De Neys, 2017; M. T. S. Raoelison et al., 2020; Thompson & Johnson, 2014) studies show that even though system 2 is hampered by memory load and time limit, conflict detection is still revealed. When system 2 was knocked out, the given response was produced by system 1.

Bago & De Neys (2017, 2019) are interested in the DI model of dual processes. According to the model, the intuitive system 1 response is given faster than the deliberative system 2 response. Bago & De Neys (2017, 2019) made several experiments to test the course assumption of the DI model. In another experiment, they used concurrent memory load and time limit to knock system 2 out. A two response paradigm was used to ensure that the initial response is to be heuristic. The initial response was given under executive load and time pressure. Then, the questions were asked to participants again without memory load and time limit. Therefore, participants had extra time to answer the questions. The syllogisms, the base rate task, and bat and ball problems were used as reasoning tasks in total twelve experiments (Bago & De Neys, 2017, 2019). The results demonstrated that initial and final incorrect responses labeled as “00” became the dominant response. This response was selected % 48 to %76 of the trials by the participants. In %7 to %10 of trials was observed that the initial response was incorrect and the final response was correct. The chosen responses labeled as “00”, and “01” are the predicted responses through the DI model. The unexpected response pattern, which is labeled as “11” both initial and final correct responses, is %15 to %42 of trials.

1.5. Belief Bias Effect and Time Pressure

The rapid response paradigm, which was initiated by Roberts and Newton (2002), was applied by Evans & Curtis-Holmes (2005) to evaluate the correlation between syllogistic reasoning and the belief bias effect. In this study, two experimental groups were formed. Time pressure was applied to one group to respond quickly whereas there was no time pressure for the other group. 12 syllogisms were served to each participant in an independently randomized order. The group that had a time limit had to select yes or no in 10 seconds. The other group has no time limit. It means that participants in this group had a chance to choose yes or no freely. The results indicated that the belief bias effect is more effective in the rapid response group than in the free time group. In the free time group, the syllogisms with valid conclusions had a higher belief bias effect than the syllogisms with invalid conclusions. It is shown that the belief bias effect is greater for the syllogisms with invalid than the syllogisms with valid. These results promote the predictions of the dual process theory. When the syllogistic reasoning problem includes conflict in which believable premises but an invalid conclusion, participants are more likely to choose to accept the conclusion (Bago & De Neys, 2017).

1.6. Two-Response Paradigm

Two response paradigms have been developed to observe the heuristic and logical response of the reasoner (Thompson et al., 2011). We want the participant's first response to be intuitive. Two-response procedure was applied in the present study. Same problems were asked twice to participant in this procedure. The challenging response deadline and memory load were presented to ensure that the participant's first response was intuitive as in the studies of Bago & De Neys (2017, 2019) and M. Raoelison & De Neys (2019). The studies revealed that the accuracy scores of first responses to the reasoning questions decreased (De Neys, 2006b; Franssens & De Neys, 2009; Oatley & Johnson-Laird, 1987). When the same question is asked for the second time, the time limit and memory load are removed. Thus, the participants are free to answer the second question by spending as much time as they want and without being under any cognitive load. Thompson & Johnson (2014) used the two-response procedure. They instructed the participants to make their initial responses intuitive but did not force the participants in that direction (Thompson & Johnson, 2014). In the following studies, (Bago & De Neys, 2017, 2019; M. T. S. Raoelison et al., 2020) adopt methodological enhancement and attempt to provide that the participant's first response could be intuitive. Because the participants might not respect the instruction and they can think deliberately during initial response stage.

Also, participants, who have high cognitive capacity, are able to give logical response when they do not be forced.

For analytical thinking to engage in any task, the necessary time and cognitive resources must be available for the effort required (Frederick, 2005; Kahneman, 2011). Bago & De Neys (2017, 2019), and Raoelison et al. (2020) applied the two response paradigm, which they strengthened methodologically. Whereas the first problem was asked with time pressure, cognitive resources were loaded with secondary load tasks synchronously. This procedure aims to inhibit cautious thinking and to reveal an intuitive response to the first question. When the same question was asked again, the time limit and cognitive load were removed. Therefore, cognitive resources might cause the correct responses by making the necessary effort.

1.7. Bat and Ball Problems

"*An apple and an orange cost $1.80 in total.*

The apple costs $1 more than the orange.

How much does the orange cost?

- *20 cents*
- *40 cents*
- *80 cents*
- *120 cents* "

This question is taken from Bago & De Neys (2019). The reason for applying modified content bat and ball problems is to eliminate familiarity with the standard bat and ball problem. Content modified bat and ball problems were used to extract such a confounding effect.

In conflict problems, the relational statement makes it difficult. For instance, in the above questions relational “more than” statement causes conflict in the participant’s mind. When the question does not include a relational statement, conflict does not occur. The heuristic system might lead to a correct response. The following item appears no conflict problem which is taken from Bago & De Neys (2019).

“*An apple and an orange cost $1.80 in total.*

The apple costs $1.

How much does the orange cost?

- *20 cents*
- *40 cents*
- *80 cents*
- *120 cents* “

Turkish version of the problem is seen as:

Bir elma ve bir portakalın toplam fiyatı 1.80 TL’dir.

Elmanın fiyatı 1 TL’dir.

Portakalın fiyatı ne kadar?

- *20 kuruş*
- *40 kuruş*
- *80 kuruş*
- *120 kuruş*

When the problem does not include a relational statement, each participant should produce the correct answer since the sample group is the university students. They learnt basic math operations. In addition to that, the heuristic system guides participants to give the correct response in no conflict questions.

1.8. Syllogistic Reasoning

Syllogisms are a sequence of statements that result in a conclusion. The conclusion is the final statement of the argument (Kenneth R.H., 2007). If we evaluate the conclusion by using rules of inference, we examine statements in terms of validity. Analytic thinking depends on validity evaluation among arguments (De Neys, 2017). A valid conclusion means that the conclusion must follow logically from the major premise and the minor premise. A reasoner should deduce the conclusion from the premises (Kenneth R.H., 2007). The validity of a conclusion is just based on the premises. The conclusion cannot be plausible in the real world. It can be also valid. The premises do not have to be true. The relationship between the premises and the conclusion was considered in the syllogisms. Therefore, believability is not the measure of validity. Rules of inference are used to arrive at valid arguments. Some systematic deviations from the rules are termed “fallacies” (Kahneman, 2011). The fallacies lead to invalid arguments. When a person used incorrect reasoning form including a fallacy, the validity evaluation becomes erroneous.

The syllogistic reasoning task was asked to the participant in this study. The syllogisms are taken from the work of Bago & De Neys (2017). Eight syllogistic reasoning problems were applied.

1.9. Effects of Memory Load in Dual Processing

De Neys (2006a) applied the conjunction fallacy problems and the card selection task with the different secondary tasks. It was one of the first dual task procedures in the dual processing of thinking research fields (De Neys, 2006a). The secondary task was used to burden working memory resources (K. E. Stanovich & West, 2003; Keith E. Stanovich & West, 2003). The analytic system requires the executive working memory resources to produce a logical answer.

Working memory is a limited capacity system. Information is stored and manipulated during ongoing activity. Baddeley's working memory model consists of three components: the phonological loop, the central executive, and the visuospatial sketch pad (Baddeley, 1992). The central executive operates attention including how the attention is divided among tasks and also how the attention is switched among tasks. In the dual task, the central executive component plays a crucial role in terms of the accuracy of a response and reaction times. De Neys (2006a) applied the secondary attention demanding task which was the finger tapping task. Therefore, the central executive component of working memory was manipulated. Participants answered the questions while they were tapping on "V, B, N, M". The executive load caused the correct response to decrease. The participants made more mistakes in the conjunction fallacy problems and the card selection task under the cognitive burden. In the card selection task, the incorrect matching card selections increased significantly. In another experiment from De Neys (2006a), the dot memory task was used to burden working memory resources. The incorrect responses increased. Results show that when the working memory resources are occupied by the load, the heuristic system becomes dominant. The responses, which are produced by system 1, were given.

1.10. Working Memory Capacity and Dual Processing in Reasoning

De Neys (2006b) tested the performance of the 308 first year psychology students, who are varying in working memory capacity, on the reasoning task. Working Memory Capacity (WMC) was measured for all participants by using the Operation Span task (La Pointe & Engle, 1990).

After measurement of WMC, participants were randomly assigned the load conditions. De Neys (2006b) used the dot memory task as concurrent memory load. There was a 3x3 matrix which is filled with three or four dots. In the high load condition, complex four dot patterns were formed in the matrix. In the low load condition, three dots on the horizontal line were formed in the matrix. The syllogisms were asked to participants. Then, the empty matrix was displayed. Participants filled the empty matrix with what they remembered from the matrix presented to them earlier. The subjects answered eight syllogistic reasoning questions, which were received from Markovits & Nantel (1989). Whereas four of the problems had conflicts, the other four problems did not have conflict. When believability and validity of the conclusion become incongruent, the questions refer to as conflict items. The other four problems did not have conflict items. The validity of the conclusion is consistent with believability of the conclusion (De Neys, 2006b).

Analytic reasoning is highly demanding on working memory resources. Therefore, the working memory capacity of a participant affects the analytic reasoning process. Individual differences in WMC are interpreted through the dual task procedure. Reasoning with secondary working memory tasks creates a conflict in the participant’s mind (Evans, 2003). It is assumed that System 1 produces the correct response when the problem does not contain conflict (Evans

& Stanovich, 2013). Even though the cognitive load is presented, the performance of participants does not affect since system 1 is not associated with computational resources. System 2 relies on logical standards, which requires executive resources. Therefore, it is supposed that participants' reasoning performance decrease under the load condition (Evans, 2003).

Dual process approach predicts that under memory load, reasoning performance regress. Because executive working memory resources were burdened by cognitive load. Activation of the analytic system demands executive resources.

The result of the study by De Neys (2006b) showed that in conflict problems incorrect heuristic responses increased. In the dual task, participants with high spans gave more correct responses than participants with low spans, when the conflict arises (De Neys, 2006b). In conflict items, the mean number of correct answers for all span groups decreased under complex dot pattern load. The concurrent memory load did not affect the performance in the reasoning task with no-conflict items. In the low span group, the load interfered with comprehension of the syllogisms in which the conclusions have a conflict between logic and belief.

System 2, which includes analytical thinking, demands executive resources along with working memory. In the designed experiment, a conflict occurs because both the working memory is loaded and the thinking for the correct answer takes place. Therefore, since more effortful thinking is inhibited by memory load, intuitive thinking, which is automatic, effortless, unconscious, associative, and does not require executive functions, plays a crucial role in producing the response (Bago & De Neys, 2017, 2019; De Neys, 2006b, 2006a; Kahneman, 2002).

1.11. Hypotheses

This study was planned to examine reasoning from the dual process framework. The generic question is as follows: How does reasoning process when biased and genuine responses were given? This research was designed to observe both intuitive and analytical thinking processes. First, we aimed at the test time course assumption of the default-interventionist model. It is expected that the heuristic responses produced by system 1 are given faster, and the responses produced by system 2 are given slower. A two response paradigm (Thompson et al., 2011) was used to observe both intuitive and deliberative responses to the problems. This paradigm is based on asking questions twice. When the problems were asked the first time, time limit, countdown from 10 seconds, and memory load as dual task were applied. Cognitive load and time pressure cause the activation of system 2 to be inhibited. Therefore, the first responses were given as heuristics labeled by system 1. When the questions include conflict items, in order to give the correct response, it is necessary to develop conflict detection sensitivity. This is the property of system 2. When the system 2 activation is blocked in the first questions by cognitive load and time limit, it is hypothesized that incorrect responses are given to the first questions which include conflict items. The incorrect responses given to the first questions, which contain conflict items, may change when the question is asked for the second time. Because when the factors that prevent system 2 are removed, if the necessary cognitive effort is made, the conflict detection sensitivity develops. In this situation, the reaction times increase. Therefore, it was another expectation that the initial incorrect and final correct responses category in the experiment questions has the highest response time. The hypothesis is that participants spend more time to respond correctly than incorrectly to experiment problems. The questions, which include no conflict item, are asked as control. The expectation in these questions is that the first and final responses are correct.

In the confidence scale scores between P1 problems and P2 problems, it is the expectation that P1 confidence scale scores are significantly lower than P2 confidence scale scores. If the evaluation is made for the direction of change categories, “00” responses and “11” responses categories have higher confidence scale scores. When the first and second responses are within the same response category such as “00” responses, the confidence scale scores increase. Thompson et al. (2011) described this situation as the “feeling of rightness”.

The originality aspect of this study is that it was conducted with a population whose native language is Turkish. Therefore, the tasks, which are taken from the studies, were translated into Turkish. These tasks have not been applied to a Turkish speaker population until now. It is aimed to observe similar results in the different languages and therefore in the different cultures, as reasoning problems allow a language-based measurement.

2. METHOD

2.1. Participants

A total of 114 undergraduate psychology students (age range: 19-25, 89 female, 25 male) at Yeditepe University, Turkey, attended the presented experiment. The subjects, who are native Turkish speakers, have no history of psychiatric disease. Participants acquired extra points in their overall Cognitive Psychology and Physiological Psychology courses grade. This research was approved by Yeditepe University Humanities and Social Research Ethics Committee.

2.2. Materials

2.2.1. Modified version of the bat and ball task

Modified version of the bat and ball problems, which are taken from the study of Bago & De Neys (2019), were served to participants to solve. There are eight problems in total. Whereas four of them are called control (no-conflict) problems, the other four problems include conflict items which are called experiment problems. There are four response options for each question. The conflict problems include a relational “more than” statement. When the relational statement was removed, the problem does not include conflict. These problems are the control problems.

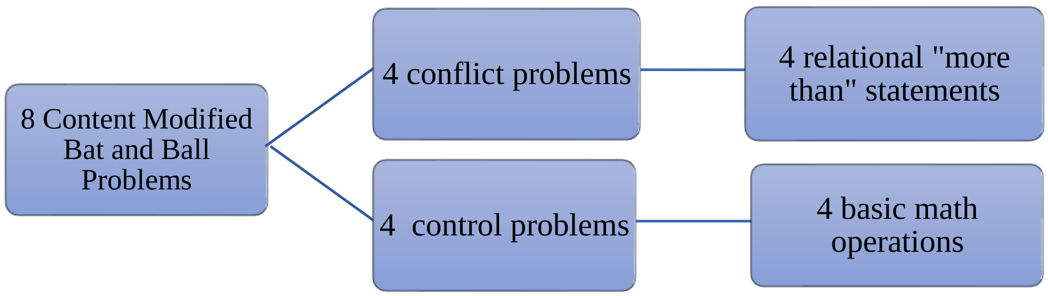

Figure 2. Distribution of experiment and control questions in the content modified bat and ball problems.

The problems were applied to a sample group whose native language is Turkish, hence, all the problems were translated into Turkish. Not only the questions were translated into Turkish, but also the units in the questions were replaced with those commonly used in Turkey. For example, the dollar sign ($) was shown as Turkish lira (TL). The following illustrates the Turkish version of the conflict problem,

"Bir elma ve bir portakalın toplam fiyatı 1.80 TL'dir.

Elmanın fiyatı portakalın fiyatından 1 TL daha fazladır.

Portakalın fiyatı ne kadar?

- *20 kuruş*
- *40 kuruş*
- *80 kuruş*
- *120 kuruş"*

The expectation was that each participant ought to respond correctly to problems that do not contain conflict. The following demonstrates the no conflict problem which is applied.

"*Bir kalem ve bir silginin toplam fiyatı 1.10 TL'dir.*

Kalemin fiyatı 1 TL'dir.

Silginin fiyatı ne kadar?

- *1 kuruş*
- *5 kuruş*
- *10 kuruş*
- *15 kuruş* "

Table 1.

Turkish version of the modified version of the bat and ball problems was used in the present experiment

Bir kalem ve bir silginin toplam fiyatı 1.10 TL'dir. Kalemin fiyatı 1 TL'dir. Silginin fiyatı ne kadar? A)1 kuruş B)5 kuruş C)10 kuruş D)15 kuruş
Bir dergi ve bir muzun toplam fiyatı 2.60 TL'dir. Derginin fiyatı muzun fiyatından 2 TL daha fazladır. Muzun fiyatı ne kadar? A)15 kuruş

B)30 kuruş C)60 kuruş D)90 kuruş
Bir peynir ve bir ekmeğin toplam fiyatı 2.90 TL’dir. Peynirin fiyatı 2 TL'dir. Ekmeğin fiyatı ne kadar? A)15 kuruş B)45 kuruş C)90 kuruş D)135kuruş
Bir elma ve bir portakalın toplam fiyatı 1.80 TL’dir. Elmanın fiyatı portakalın fiyatından 1 TL daha fazladır. Portakalın fiyatı ne kadar? A)20 kuruş B)40 kuruş C)80 kuruş D)120 kuruş
Bir sandviç ve bir sodanın toplam fiyatı 2.50 TL’dir. Sandviçin fiyatı 2 TL’dir. Sodanın fiyatı ne kadar? A)5 kuruş B)25 kuruş

C)50 kuruş D)75 kuruş
Bir şapka ve bir kurdelenin toplam fiyatı 4.20 TL'dir. Şapkanın fiyatı kurdelenin fiyatından 4 TL daha fazladır. Kurdelenin fiyatı ne kadar? A)5 kuruş B)10 kuruş C)20 kuruş D)30 kuruş
Bir kahve ve bir kurabiyenin toplam fiyatı 2.40 TL'dir. Kahvenin fiyatı 2 TL'dir. Kurabiyenin fiyatı ne kadar? A)10 kuruş B)20 kuruş C)40 kuruş D)60 kuruş
Bir kitap ve bir ayracın toplam fiyatı 3.30 TL'dir. Kitabın fiyatı ayracın fiyatından 3 TL daha fazladır. Ayracın fiyatı ne kadar? A)5 kuruş B)15 kuruş

C)30 kuruş D)45 kuruş

2.2.2. Syllogistic reasoning task

Syllogistic reasoning problems were taken from the study which is made by Bago & De Neys (2017). The problems were given to the participants whose native language is Turkish, hence, all the problems were translated into Turkish. The one problem was changed because the meaning of the problem is not clear when translated into Turkish. Eight syllogistic reasoning problems were given to participants. Whereas four of them include conflict items, the other four do not contain conflict. All problems consisted of a major premise, a minor premise, and a conclusion. The question “Does the conclusion follow logically?” was seen under premises and a conclusion. The options “Yes and No” were placed under the question. Participants were able to choose yes or no.

In this task, the validity of the conclusion depends on deductive logic rules. When the conclusion follows the premises by the law of deductive logic, the conclusion is evaluated in terms of validity. The believability of the conclusion relies on beliefs. Beliefs cause participants’ response tends to be automatic and intuitive. Therefore, conflict arises when the believability and validity of a conclusion become incongruent. There were two types of incongruent items including an unbelievable-valid conclusion and a believable-invalid conclusion. Two problems with an unbelievable-valid conclusion and two problems with a believable-invalid conclusion were used to counterbalance purposes. When the logical validity and believability of a conclusion become congruent, conflict disappears. Two types of congruent items were used including a believable-valid conclusion and an unbelievable-invalid conclusion.

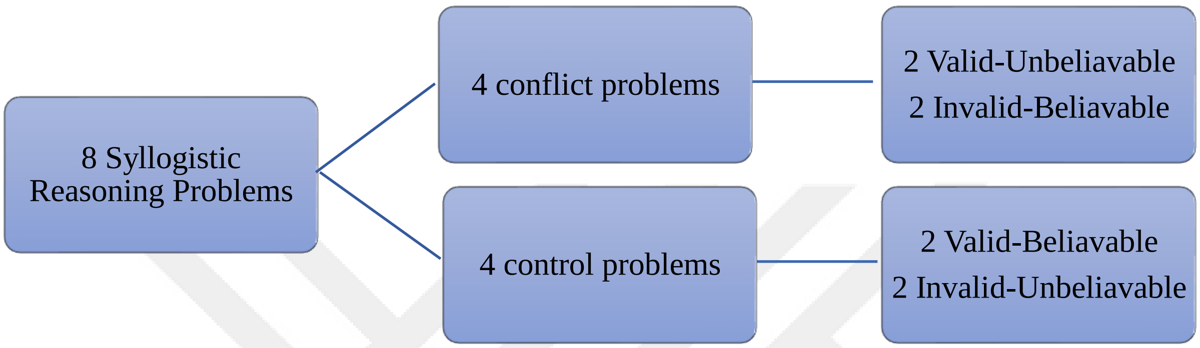

Figure 3. Distribution of experiment and control questions in the syllogistic reasoning problems

Two problems with a believable-valid conclusion and two problems with an unbelievable-invalid conclusion were used to counterbalance purposes. The same content was used for both conflict problems and no-conflict, control problems. The difference between conflict and no-conflict problems is the order in which the minor premise and the conclusion.

Table 2.

Turkish version of the syllogistic reasoning problems was used in the present experiment

Tüm çiçekler ışığa ihtiyaç duyar Güller ışığa ihtiyaç duyar Sonuç: Güller çiçektir Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	Conflict: Invalid-Believable

Tüm pilli cihazların şarjı biter Telefonlar pilli cihazlardır Sonuç: Telefonların şarjı biter Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	No Conflict: Valid-Believable
Tüm memeli hayvanlar yürüyebilir Balinalar memelidir Sonuç: Balinalar yürüyebilir Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	Conflict: Valid-Unbelievable
Tüm taşıtların tekerlekleri vardır Bavulların tekerlekleri vardır Sonuç: Bavullar taşıttır Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	No Conflict: Invalid-Unbelievable
Tüm kuşlar kanatlıdır Kargalar kanatlıdır Sonuç: Kargalar kuştur Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	Conflict: Invalid-Believable
Tüm köpeklerin burnu vardır Labradorlar köpektir	No Conflict: Valid-Believable

Sonuç: Labradorların burnu vardır Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	
Tüm çiçekli bitkilerin yaprakları vardır Kaktüsler çiçekli bitkilerdir Sonuç: Kaktüslerin yaprakları vardır Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	Conflict: Valid-Unbelievable
Tüm silahlar mermi ateşler Su tabancaları mermi ateşler Sonuç: Su tabancaları silahtır Sonuca mantıksal olarak varılıyor mu? Evet/Hayır	No Conflict: Invalid-Unbelievable

2.3. Procedure

The experiment was programmed in PsychoPy which is a psychology experiment software. The experiment run on a Dell laptop. The experiment was conducted at the Yeditepe Cognitive Psychology Laboratory. A Dell laptop, which has a 360x240 mm screen, was used. The distance of the participants from the screen was 600mm. Participants' responses and reaction times were recorded.

Participants were instructed about the response format and the reasoning tasks which are used in this experiment. Two problems were applied within the practice session in which participants received feedback. Memory load, which consists of 4 digits, was seen for 4 seconds. Then, the question, which is in the format of a syllogistic reasoning problem, was asked. The problem consisted of a major premise (e.g., “All fruits are red”), a minor premise (e.g., “Cherries are red”), and a conclusion (e.g., Cherries are fruits”). The question is “Does the conclusion follow logically?”. Yes and no are the response options to the question. The problem remained on the screen for 10 seconds. Countdown from 10 seconds was seen on the screen. Therefore, participants were able to see how much time they had left to answer the question. Participants received feedback on whether they answered the question correctly or incorrectly. A confidence scale between 1 and 5 was displayed. Participants marked how confident they were in their answers on the scale. After that, participants were asked to type the digits they could remember. They received feedback on whether they remembered the memory load or not. The same problem was asked again without concurrent memory load and challenging response deadline. Participants again received feedback on whether they answered the question correctly or incorrectly. Then the confidence scale was shown to participants for their marking once again.

The other question type, which is in the format of the bat and ball problem, was asked in the practice session. The question was “A milk and a tea cost 20 TL in total. A tea costs 5 TL. How much does the milk cost?”. There were four response options. As in the first practice question, the memory load consisting of 4 digits had been shown to the participant for 4 seconds before the problem appeared. Then the question was displayed for 10 seconds. Countdown from 10 seconds was seen on the screen. Participants received feedback on whether they answered the question correctly or incorrectly. A confidence scale between 1 and 5 was displayed. Participants marked how confident they were in their answers on the scale. After that, participants were asked to type the digits they could remember. They received feedback on whether they could remember the memory load or not. The same problem was asked again without concurrent memory load and challenging response deadline. Participants again received feedback on whether they answered the question correctly or incorrectly. Then the confidence scale was shown to participants for their marking. The instruction displayed that the practice session was over.

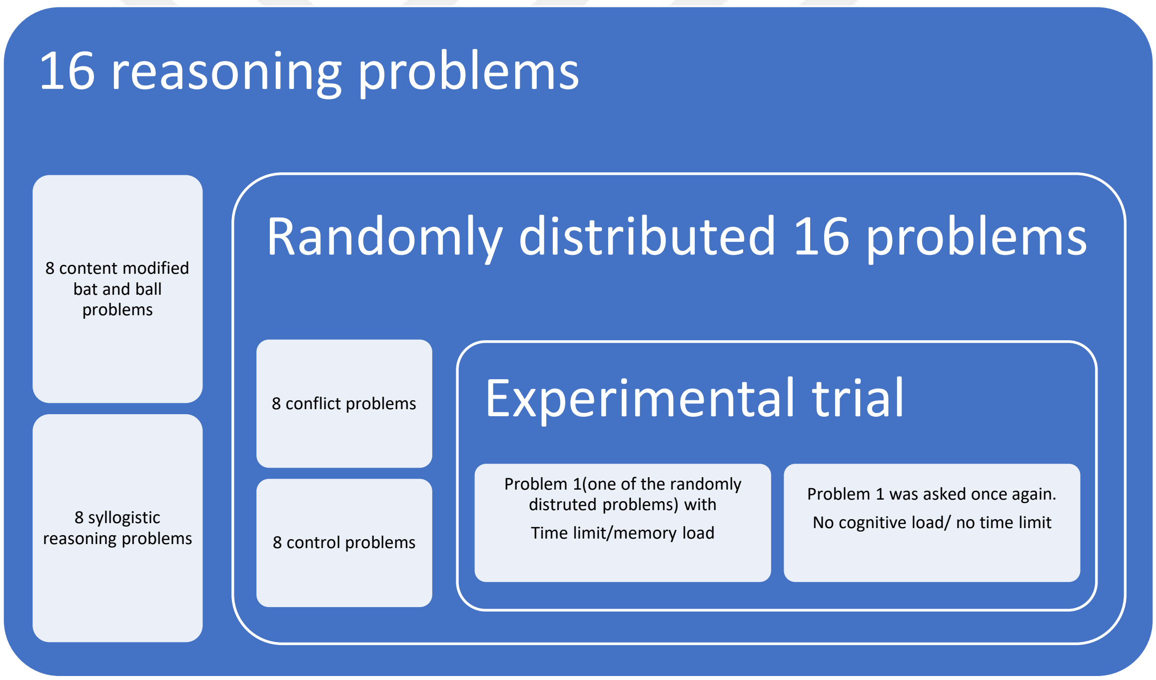

Figure 4. This figure demonstrates the question types that were used in the experiment.

Sixteen reasoning problems were randomly distributed to each participant. Therefore, there is no fixed order among the questions. This method was used to eliminate the effects of the fixed order of the questions. A trial follows that the problem was selected randomly from the set of 16 problems. This problem was asked with memory load and time limit. Then the same problem once again without memory load and time limit.

In the experiment session, sixteen reasoning problems, which consisted of syllogistic reasoning and the modified version of the bat and ball problems, were presented in a random order for each participant. In both tasks, there are 4 conflict and 4 control problems. The trials in the experiment session followed the same procedure as the practice trials. The only difference from the practice trials is that no feedback was given after answering the questions. Memory load and response deadline were applied in the first questions. The participants chose one option out of four options in the modified version of the bat and ball problems. Yes and no are the response options in the syllogistic reasoning problems. All questions were presented to each participant in random order. After the first problem was solved, the question, “How confident are you in your answer?” was asked to participants to observe the confidence of their first responses. The scale was formed from 1 to 5. While “1” demonstrates not being confident, “5” signifies confident. The same reasoning question was presented again without both time pressure and memory load. After that, the question, “How confident are you in your answer?” was asked to participants once again to detect their confidence in their final answer.

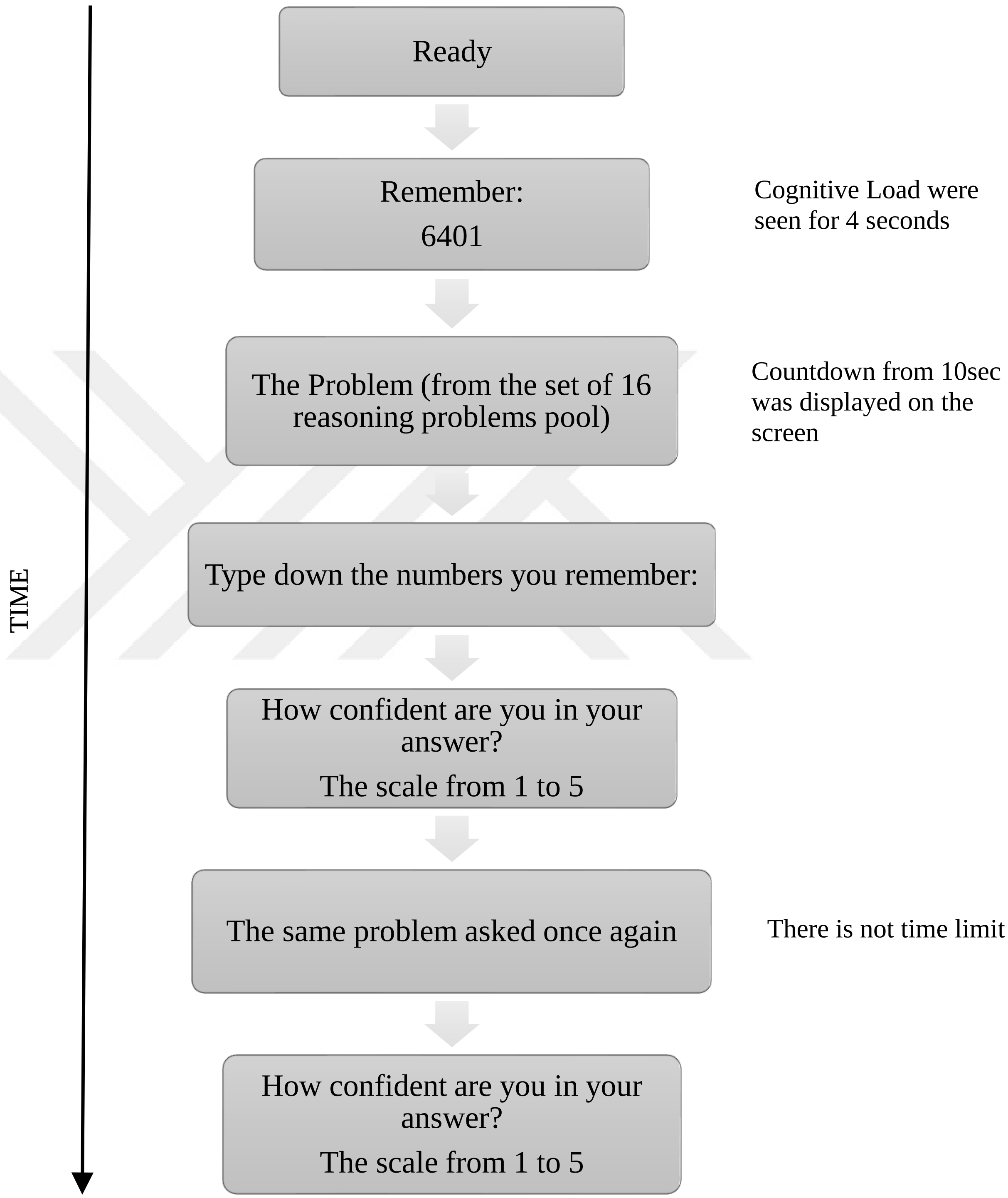

Figure 5. The demonstration of the sequence of a trial

2.3.1. Experimental Design

In the present study, the reasoning process was studied upon the modified version of the bat and ball problems and syllogistic reasoning problems by using two response paradigm and direction of change analysis. We replicated the typical direction of change analysis of conflict detection with the two response paradigm.

Eight modified version of the bat and ball problems and eight syllogistic reasoning problems were used in this experiment. In both tasks, half of the problems include conflict, and the rest of the problems have no conflict. No conflict problems are the control problems (De Neys & Pennycook, 2019). A two response paradigm was formed. Participants respond the first answer under time pressure and cognitive load. Participants are expected to respond as quickly as possible because the first, heuristic response comes directly from the mind (Bago & De Neys, 2017; M. T. S. Raoelison et al., 2020). Then, the same problem is presented again to the participants. The participants have as much time as they want to think and this time there is no cognitive load. Whereas the incorrect answer is labeled “0”, the correct answer is labeled “1”. For instance, the participant’s first answer is wrong and the final answer is correct. It is demonstrated as “01”. “00” means that both the first response and final response are incorrect. and “11” means that both the first response and final response are correct. “01” signifies that first response is incorrect and final response is correct. “10” denotes that first response is correct and final response is incorrect.

3. RESULTS

The questions which are presented with concurrent load and time pressure are named P1. The questions, which are presented without cognitive load and time pressure, are named P2. First descriptive statistics were made to assess participants' performance in the reasoning tasks.

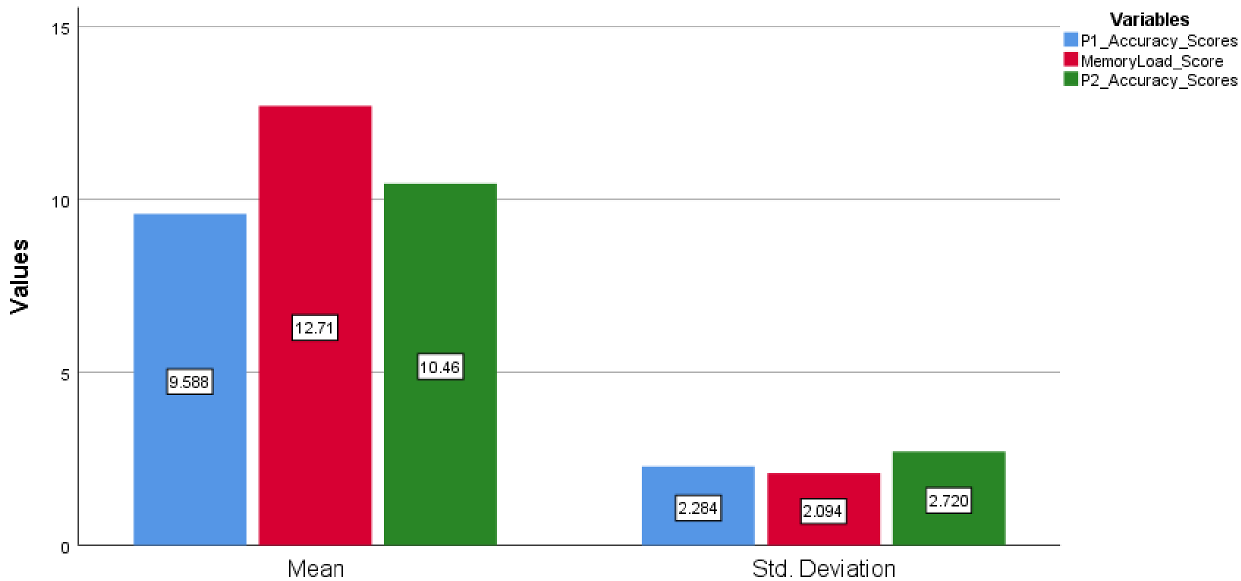

Figure 6. Mean and standard deviation values of accuracy scores of P1, memory load, and P2.

The mean of the P2 accuracy scores (*M* = 10.46, *SD* = 2.720) is more than the mean of the P1 (*M* = 9.588, *SD* = 2.284) accuracy scores. The mean value of memory load scores is 12.71. The standard deviation of the memory load scores is 2.094.

When the normality analyzes were performed on the data for each variable including accuracy scores, reaction time scores, memory load scores, and confidence scale scores, it was observed that it did not fit in the normal distribution. Therefore, non-parametric analyzes were applied to observe the differences.

3.1. Accuracy Scores Differences Based on Problem Types

Accuracy scores were compared between P1 and P2. A Wilcoxon signed rank test was used since P1 accuracy scores and P2 accuracy scores are related and non-parametric variables. A significant difference was observed between P1 accuracy scores (*Mdn* = 9.00, *n* = 114) and P2 accuracy scores (*Mdn* = 10.00, *n* = 114), *z* = -5.916, *p* = .000 with a moderate effect size, *r* = 0.39. A Wilcoxon signed rank test revealed that P2 correct response scores were significantly higher after removal of memory load and time pressure compared to P1 correct response scores.

3.2. Reaction Times Differences Based on Problem Types

P1 reaction time scores and P2 reaction time scores were recorded. A Wilcoxon signed rank test was used since P1 reaction time scores and P2 reaction time scores are related and non-parametric variables. There was a significant difference between P1 reaction time scores (*Mdn* = 87.75, *n* = 114) and P2 reaction time scores (*Mdn* = 107.57, *n* = 114), *z* = -6.722, *p* = .000 with a moderate effect size, *r* = 0.44. Therefore, P1 reaction time scores are significantly lower than P2 reaction time scores.

3.2.1. Mean Reaction Times Difference Across Correct Responses Based on Problem Types

Reaction times of the correct responses were selected from the data and the mean values of reaction times were calculated for both P1 questions and P2 questions. A Wilcoxon signed rank test was applied since mean values of the correct responses in P2 reaction times are non-parametric. The test revealed that there was a significant difference in mean values of reaction times of the correct responses between P1 questions (*Mdn* = 5.524, *n* = 114) and P2 questions (*Mdn* = 6.456, *n* = 114), *z* = -5.130, *p* = .000 with a moderate effect size, *r* = 0.34. Mean values of P1 correct response's reaction times are significantly lower than mean values of P2 correct response's reaction times.

3.2.2. Correlation Between P1 Accuracies and P1 Reaction Times

Spearman's rank correlation coefficient analysis was applied between P1 reaction time scores and P1 accuracy scores to questions. This analysis reveals that there is a positive and significant correlation between P1 reaction time scores and P1 accuracy scores, $r_{spearman} = .260$, $p < .001$.

Table 3.

Spearman's Rank Correlation Coefficient analysis applied between P1 reaction time scores and P1 accuracy scores

Variable	*n*	*M*	*SD*	*1*	*2*
1. P1 Accuracy Scores	114	9.59	2.284	-	.260**
2. P1 Reaction Time Scores	114	86.03	15.21		-

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

This correlation is significant at the 0.01 level (2-tailed). When the number of correct response increase, the reaction times extend in the first presentation of the problems.

3.2.3. Correlation Between P2 Accuracies and P2 Reaction Times

Spearman's Rank Correlation Coefficient analysis was used between P2 reaction time scores and P2 accuracy scores. It was indicated that there is a positive and significant correlation between P2 reaction time scores and P2 accuracy scores, $r_{\text{spearman}} = .453$, $p < .001$.

Table 4.

Spearman's Rank Correlation Coefficient analysis applied between P2 reaction time scores and P2 accuracy scores

Variable	*n*	*M*	*SD*	*1*	*2*
1. P2 Accuracy Scores	114	10.46	2.720	-	.453**
2. P2 Reaction Time Scores	114	123.2	62.68		-

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

A positive correlation was observed between P2 accuracy scores and P2 reaction time scores and this correlation is significant at the 0.01 level (2-tailed).

3.2.4. Correlation Between P1 Accuracies and Mean Values of P1 Correct Responses' Reaction Times

Spearman's rank correlation coefficient analysis was applied between mean values of P1 correct responses' reaction times and P1 accuracy scores to questions. This analysis appeared that there is a positive and significant correlation between mean values of P1 correct responses' reaction times and P1 accuracy scores, $r_{\text{spearman}} = .351$, $p < .001$.

Table 5.

Spearman's Rank Correlation Coefficient analysis applied between the mean value of P1 correct responses' reaction times and P1 accuracy scores

Variable	*n*	*M*	*SD*	*1*	*2*
1. P1 Accuracy Scores	114	9.59	2.284	-	.351**
2. Mean Value of P1 Reaction Time Scores	114	5.454	0.999		-

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

This correlation is significant at the 0.01 level (2-tailed). When the number of correct response increase, the mean values of correct responses' reaction times extend in the first presentation of the problems.

3.2.5. Correlation Between P2 Accuracies and Mean Values of P2 Correct Responses' Reaction Times

Spearman's Rank Correlation Coefficient analysis was used between mean values of P2 correct responses' reaction times and P2 accuracy scores. It was appeared that there is a positive and significant correlation between mean values of P2 correct response's reaction times and P2 accuracy scores, $r_{spearman} = .591$, $p < .001$.

Table 6.

Spearman's Rank Correlation Coefficient analysis applied between mean values of P2 correct responses' reaction times and P2 accuracy scores

Variable	*n*	*M*	*SD*	*1*	*2*
1. P2 Accuracy Scores	114	10.46	2.720	-	.591**
2. P2 Reaction Time Scores	114	7.102	3.376		-

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

A positive correlation was observed at the 0.01 level (2-tailed). When the correct response increased in P2 questions, reaction times of correct responses increased in P2 questions.

3.3. Confidence Scale Score Difference Based on Problem Types

P1 confidence scale scores and P2 confidence scale scores were recorded. A Wilcoxon signed rank test was used to examine the difference in the confidence scale scores between P1 and P2. The test revealed that there was a significant difference between P1 confidence scale scores (*Mdn* = 68, *n* = 114) and P2 confidence scale scores (*Mdn* = 77, *n* = 114), $z = -8.838$, $p = .000$ with a strong effect size, $r = 0.58$. Therefore, P1 confidence scale scores are significantly lower than P2 confidence scale scores.

3.4. Effects of Memory Performance on the P1 Responses

The mean value of the P1 accuracy score is 9.588 and the mean value of the memory load score is 12.71.

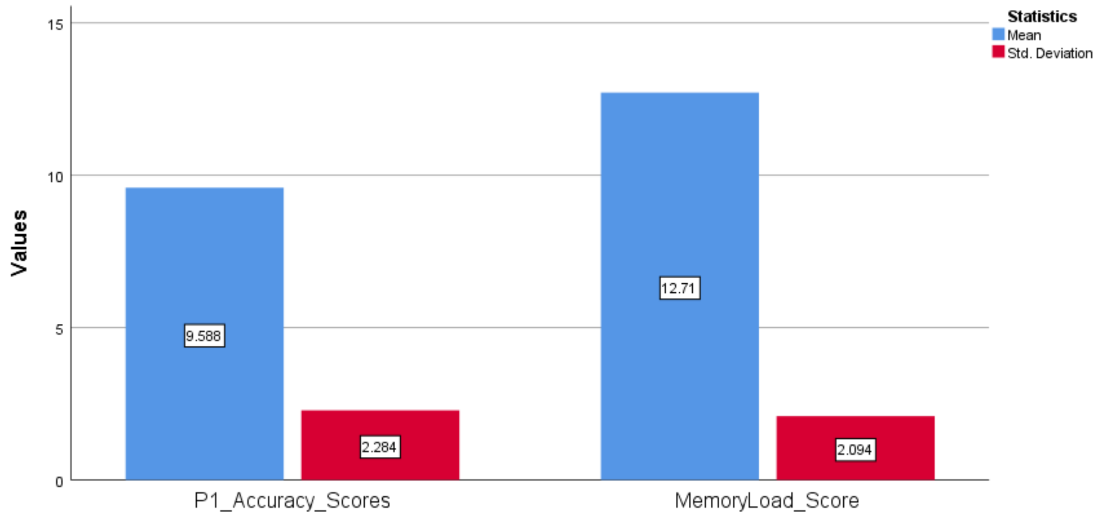

Figure 7. Mean and standard deviation values of accuracy scores of P1 and memory load scores.

It was examined whether the answers given to the first questions changed with whether the numbers are remembered or not.

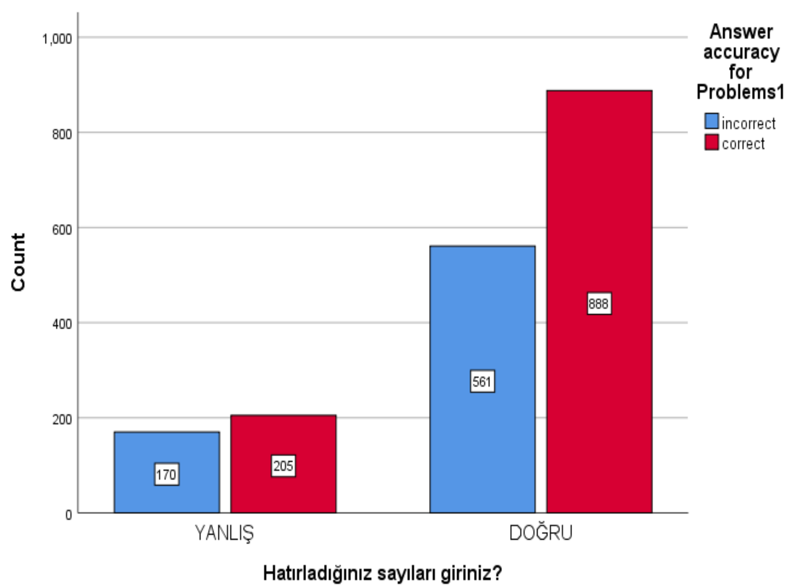

Figure 8. Number of correct and incorrect responses across memory performance

The chi-square analysis was applied to observe the difference between whether the memory load was remembered correctly or not and the correct and incorrect answers to the first questions.

Table 7.

Correct and incorrect responses difference based on memory load condition

		Answer Accuracy for Problems 1					
		Correct		*Incorrect*		x^2	*p*
		Frequency	*Percent(%)*	*Frequency*	*Percent(%)*		
Hatırladığınız sayıları giriniz	Correct	888	61.3	561	38.7	5.431	.020
	Incorrect	205	54.7	170	45.3		

When the memory load is remembered correctly in the P1 questions, %61.3 of the responses were answered correctly to the reasoning questions, and %38.7 of the responses were answered incorrectly to the reasoning questions. When memory load is not remembered in the P1 questions, %54.7 of the responses were given correctly and %45.3 of the responses were given incorrectly. According to the Chi-square test, the remembering memory load difference between correct and incorrect answers was significant (p=0.020, p<0.05) where the percentage of the correct responses (%54.7) in the not remembering the memory load was more than the percentage of the incorrect response (38.7) in the remembering of memory load correctly. It has been found that the correct answer is given more commonly in the case of not remembering the memory load correctly. Therefore, the frequency of correct response and correct remembering of the cognitive load has the highest frequency value (See Table 5).

3.4.1. Correlation Between P1 Accuracy and Memory Load Score

Spearman’s Rank Correlation Coefficient analysis was used to investigate whether there was a correlation between the accuracy scores of the memory load presented simultaneously in the first questions and the accuracy of the answers to the questions. It was revealed that there is a negative and nonsignificant correlation between P1 accuracy scores and memory load scores, $r_{spearman} = -.013$, $p = .892$.

Table 8.

Spearman’s Rank Correlation Coefficient analysis applied between P1 accuracy scores and memory load scores

Variable	*n*	*M*	*SD*	*1*	*2*
1. P1 Accuracy Scores	114	9.588	2.284	-	-.013
2. Memory Load Scores	114	12.71	2.094		-

$^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

There is a negative correlation between P1 accuracy scores and memory load scores but this correlation is not significant. When the number of correct response increase in the P1, correct recall of memory load decreases.

3.5. Accuracy Difference of Syllogistic Control (No Conflict) Questions

A Wilcoxon Signed Rank Test revealed that accuracy scores difference were not significant between P1 syllogisms control questions scores (*Mdn* = 3.00, *n* = 114) and P2 syllogisms control questions scores (*Mdn* = 3.00, *n* = 114), *z* = -1.35, *p* = .178. Therefore, there is not a significant difference between these groups.

3.5.1. Reaction Time Difference of Syllogistic Control (No Conflict) Questions

A Wilcoxon Signed Rank Test showed that there was not a significant difference between reaction time scores of the P1 control syllogisms (*Mdn* = 22.74, *n* = 114) and reaction time scores of the P2 control syllogisms (*Mdn* = 23.09, *n* = 114), *z* = -1.86, *p* = .063. Hence, reactions times scores were not significantly different in P1 and P2 control syllogisms.

3.6. Accuracy Difference of Syllogistic Experiment (Conflict) Questions

A Wilcoxon Signed Rank Test showed that accuracy scores difference was significant between P1 syllogisms experiment (Conflict) questions scores (*Mdn* = 1.00, *n* = 114) and P2 syllogisms experiment (Conflict) questions scores (*Mdn* = 2.00, *n* = 114), *z* = -2.725, *p* = .006 with a small effect size, *r* = 0.18. Therefore, there are significant differences in syllogistic experiment (Conflict) questions between P1 and P2. The correct response scores were significantly higher after memory load and time limit were removed compared to the correct response scores of syllogistic experiment (Conflict) questions with concurrent memory load and time pressure.

3.6.1. Reaction Time Difference of Syllogistic Experiment (Conflict) Questions

The reaction time difference among experiment (conflict) syllogisms was observed by using a Wilcoxon Signed Rank Test which revealed that there is a significant difference between

P1 experiment (conflict) syllogisms (*Mdn* = 22.66, *n* = 114) and P2 experiment (conflict) syllogisms (*Mdn* = 26.13, *n* = 114), *z* = -4.282, *p* = .000 with a small effect size, *r* = 0.28. The reaction time of P2 experiment syllogisms is significantly higher than the reaction time of P1 experiment syllogisms.

3.7. Accuracy Difference of Modified Version of the Bat and Ball Control (No Conflict) Questions

A Wilcoxon Signed Rank Test revealed that accuracy scores difference was significant between P1 bat and ball control (No Conflict) questions scores (*Mdn* = 4.00, *n* = 114) and P2 bat and ball control (No Conflict) questions scores (*Mdn* = 4.00, *n* = 114), *z* = -3.153, *p* = .002 with a small effect size, *r* = 0.21. The correct response scores were significantly higher after memory load and time limit were removed compared to the correct response scores of bat and ball control (No Conflict) questions with concurrent memory load and time pressure.

3.7.1. Reaction Time Difference of Modified Version of the Bat and Ball Control (No Conflict) Questions

A Wilcoxon Signed Rank Test showed that there was a significant difference between reaction time scores of the P1 control bat and ball questions (*Mdn* = 19.29, *n* = 114) and reaction time scores of the P2 control bat and ball questions (*Mdn* = 19.45, *n* = 114), *z* = -2.162, *p* = .031 with a small effect size, *r* = 0.14. The reaction time of P2 control bat and ball questions is significantly higher than the reaction time of P2 control bat and ball questions.

3.8. Accuracy Difference of Modified Version of the Bat and Ball Experiment (Conflict) Questions

A Wilcoxon Signed Rank Test indicated that accuracy scores difference was significant between P1 bat and ball experiment (Conflict) questions scores (*Mdn* = 0.00, *n* = 114) and P2 bat and ball experiment (Conflict) questions scores (*Mdn* = 0.00, *n* = 114), *z* = -5.199, *p* = .000 with a moderate effect size, *r* = 0.34. The correct response scores were significantly higher after memory load and time limit were removed compared to the correct response scores of bat and ball experiment (Conflict) questions with concurrent memory load and time pressure.

3.8.1. Reaction Time Difference of Modified Version of the Bat and Ball Experiment (Conflict) Questions

The reaction time difference between experiment (conflict) bat and ball questions was observed by using a Wilcoxon Signed Rank Test which indicated that there is a significant difference between P1 experiment (conflict) bat and ball questions (*Mdn* = 21.80, *n* = 114) and P2 experiment (conflict) bat and ball questions (*Mdn* = 34.87, *n* = 114), *z* = -7.669, *p* = .000 with a strong effect size, *r* = 0.51. The reaction time of P2 experiment bat and ball questions is significantly higher than the reaction time of P1 experiment bat and ball questions.

3.9. Direction of Change Categories

3.9.1. Percentages of Trials and Frequencies

In the control syllogistic reasoning problems between P1 and P2, there is no significant difference. In most trials (72.4%), participants responded correctly to both the P1 control syllogisms and the P2 control syllogisms.

Table 9.

Percentage of trials (frequencies) within the direction of change categories for experiment and control questions

Task	Question Type	*Direction of change*			
		11	*00*	*10*	*01*
Syllogism	Control (No Conflict)	72.4% (330)	12.9% (59)	6.1% (28)	8.6% (39)
	Experiment (Conflict)	34% (155)	52% (237)	4.4 % (20)	9.6% (44)
Bat and Ball	Control (No Conflict)	94.5% (431)	2% (9)	0.2% (1)	3.3% (15)
	Experiment (Conflict)	25.7% (117)	58.6% (267)	2.4 (11)	13.4% (61)

11 = response 1 correct; response 2 correct, 00 = response 1 incorrect; response 2 incorrect, 10 = response 1 correct; response 2 incorrect; 01 = response 1 incorrect; response 2 correct

In the bat and ball control questions, %94.5 (431) trials were responded correctly. When problems contained conflict items, the number of correct responses decreased (See Table 9). Whereas %34 trials were responded correctly to experiment syllogisms, %25.7 trials were responded correctly to experiment bat and ball problems.

3.9.2. Mean of Reaction Times

The mean values of reaction times were calculated within the direction of change categories for experiment and control problems. Significant differences were observed in both experiment syllogisms and experiment bat and ball problems between P1 and P2. When the initial responses were changed during the final response in the experiment problems, the reaction times were dramatically increased.

Table 10.

Mean value of reaction times within direction of change categories for experiment and control questions

Task	Question Type	*Direction of change*			
		11	*00*	*10*	*01*
Syllogism	Control (No Conflict)	11.1956	12.9086	19.6664	14.9223
	Experiment (Conflict)	14.3817	10.4928	20.3976	19.1723
Bat and Ball	Control (No Conflict)	9.8590	10.7182	7.6176	18.1576
	Experiment (Conflict)	16.6046	14.2263	25.3359	30.3633

11 = response 1 correct; response 2 correct, 00 = response 1 incorrect; response 2 incorrect, 10 = response 1 correct; response 2 incorrect; 01 = response 1 incorrect; response 2 correct

3.9.3. Mean of Confidence Scale Scores

The mean values of confidence scale scores were calculated within the direction of change categories for experiment and control questions. When the responses were not changed including “00” responses and “11” responses, the mean of confidence scale scores have higher values compared to the response categories which changed their initial response during the final response.

Table 11.

Mean value of confidence scale scores within the direction of change categories for experiment and control questions

Task	Question Type	*Direction of change*			
		11	*00*	*10*	*01*
Syllogism	Control (No Conflict)	8.56	8.08	7.21	7.67
	Experiment (Conflict)	8.66	8.89	7.40	7.41
Bat and Ball	Control (No Conflict)	9.49	9.67	10	8.33
	Experiment (Conflict)	8.93	8.81	7.09	8.05

11 = response 1 correct; response 2 correct, 00 = response 1 incorrect; response 2 incorrect, 10 = response 1 correct; response 2 incorrect; 01 = response 1 incorrect; response 2 correct

4. CONCLUSION

In this study, we aimed to observe how system 1 and system 2 affect the performance in the reasoning tasks including conflict and no-conflict items for both with constraints and without constraints. The syllogistic reasoning task and the modified version of the bat and ball problems, which involved experiment and control questions, were applied. Firstly, assessments were made on accuracy scores, response time scores, and confidence scale scores between initial and final responses. It was observed that the final response scores including accuracy scores, response time scores, and confidence scale scores are higher than the initial response scores. Positive correlations between accuracy scores and reaction time scores for both the initial response stage and final response stage were observed. Then, the reaction times of the correct responses were selected. It was shown that there is also a positive correlation between the number of correct responses and the reaction times of the correct responses for both the initial and final responses stages. Therefore, these result supports the DI model assumption.

In both the syllogisms control problems and the bat and ball control questions, the predominant category is first correct and final correct responses. Thus, our hypothesis for control questions was supported.

On the other hand, it was observed that the number of correct responses decreased significantly in the both syllogisms experiment questions and bat and ball experiment questions. In the experiment problems, the main response category was the initial incorrect response and the final incorrect response (“00” responses). These results are consistent with DI dual process predictions. The biased reasoners do not detect conflicts. Therefore, the responses, which were produced by system 1, were given. In addition, these biased reasoners, who gave incorrect answers to both of the experiment questions, had higher confidence scale scores. The biased reasoners do not doubt their answers since they are not aware of the conflict.

When reaction times were observed, the results revealed that the participants, who developed sensitivity to conflict items in the experiment problems, had higher response times. If the problem contains conflict, analytic thinking gets involved where the correct answer is given (De Neys, 2006b). The conflict detection sensitivity was the characteristic of system 2. Hence, the result validates that system 2 based thinking is time consuming.

5. DISCUSSION

This experiment was designed in order to observe the properties of the heuristic response and analytic response during reasoning from a dual process perspective. It was the hypothesis that, If the heuristics, system 1, is the reason for the incorrect responses, the prevalence of incorrect responses in conflict problems would raise in the initial response stage because system 2 was inhibited by cognitive load and time pressure. Afterward, the problem is asked again without both memory load and time pressure. It means that there is no suppressor factor to system 2. Therefore, participants may evaluate the problem in terms of logical reasoning. The hypothesis was based on that system 2 needs working memory resources. When the memory load was given as a dual task, system 2's access to working memory resources was restricted. Also, the time limit creates pressure on system 2. Because system 2 needs enough time to produce a response. Evans & Curtis-Holmes (2005) revealed that the belief bias effect increased in syllogisms when the 10 seconds time limit was applied.

The results of this experiment have revealed that working memory load and time pressure caused the incorrect responses to increase. Because there is a significant difference between P1 accuracy scores and P2 accuracy scores. When memory load and time limit were removed, the correct responses increased in the final presentation of the problems (P2). Time limit and memory load made it difficult for system 2 to engage. Participants respond more correct answers to P2 questions than to P1 questions. This result is consistent with the default interventionist model. Time limit and cognitive load caused to produce heuristic response generated by system 1.

Reaction times' difference was also observed between P1 questions and P2 questions. When the time limit was removed, the reaction times of the P2 questions got longer. Also, P2 correct responses were higher than P1 correct responses. When the reaction times increased, the number of correct responses increased. In the P2 questions, participants have developed conflict detection sensitivity. Some of the biased reasoners changed their first responses, which were chosen in the P1 questions. It means that some of the participants became sensitive to conflict items in the second presentation of the problems. The increased response times also signified this conclusion. Mean reaction times difference across correct responses between P1 and P2 questions were also examined. A significant difference was observed. The mean value of correct responses' reaction times in P2 questions is significantly higher than the mean value of correct responses' reaction times in P1 questions. Correlation analysis was made to observe the relation between P1 correct responses and P1 reaction times. Significant and positive correlation were observed. When the correct responses increased, the reaction times got longer in P1 problems. The correlation between P2 correct responses and P2 reaction times was researched. The correlation was significant and positive. When the correct responses increased, the reaction times got longer in P2 problems.

The syllogistic reasoning task and the modified version of the bat and ball task were served in this research. There were experiment (conflict) and control (no conflict) questions in both tasks. The expectation was that in the control questions, both first and second responses are to be correct. DI theory supports that heuristic system 1 processing provides to generate correct responses on control (no conflict) problems in both P1 and P2 (Bago & De Neys, 2017). If the problem does not include conflict, the heuristic system produces the correct response quickly. In this case, the heuristic system likely guides the participant to select appropriate options. The guidance depends on prior semantic knowledge, which leads to "contextualizing" a problem. The knowledge as a prototype regulates the participant's judgment. Even though the predictor is not aware of what exactly he or she chooses, the heuristic system allows the participant to respond fast with less cognitive effort (De Neys, 2006b).

The vast majority of trials were answered correctly to both the P1 control bat and ball problems and the P2 control bat and ball problems and the same result was observed for P1 control syllogisms and P2 control syllogisms. Correct response scores (11, first and final correct responses) in both P1 and P2 decreased dramatically when the problems include conflict items for both syllogistic reasoning experiment questions and bat and ball experiment problems. Even though the both first and second correct responses decreased, "11" responses frequency was still high. Whereas %34 (155) trials were responded as "11" responses in the experiment syllogisms, %25.7 (117) trials were responded as "11" responses in the experiment bat and ball problems. A major challenge for the Default interventionist model is the high frequency of "11" responses in the P1 and P2 experiment problems (Bago & De Neys, 2017, 2019; M. T. S. Raoelison et al., 2020; Šrol & De Neys, 2021). There can be several reasons for "11" responses in the experiment problems, which include conflict items. The first reason might be individual differences in the mental activities among participants. Challenging response deadline and secondary cognitive load may not make difficult to the tasks enough for some participants. Because working memory capacity of some participants might be higher than the other participants. If the working memory resources are not sufficiently occupied by the cognitive load, system 2 provides to access to the working memory resources and ensures that the correct response is produced in the P1 as well. Another reason for "11" responses in the experiment problems might be random guessing. A dual task as memory load and strict deadline during the first presentation of the problems made it difficult to respond. The tasks might be too challenging for some participants. Therefore, the first presentation of experiment questions may be responded as random guessing which might be correct. Because syllogistic reasoning problems include two options. Therefore, there is a ½ (%50) probability to respond correctly. Bat and ball questions involve four choices. Hence, there is a ¼ (%25) probability to give a correct answer. Briefly, initial responses could be randomly selected.

The dominant response category consisted of the first incorrect response and final incorrect response for both reasoning experiment problems which involve conflict items. This result is consistent with the result of Bago & De Neys (2017). A significant difference was shown in the result section between P1 experiment syllogisms and P2 experiment syllogisms. %9.6 (44) trials were observed as incorrect first response and correct final response. In addition to that, a significant difference was examined between the P1 experiment bat and ball problems and the P2 experiment bat and ball problems. %13.4 (61) trials were responded as first erroneous response and final correct response. In these cases, the first heuristic incorrect response was given by biased reasoners. After the first quick response, the reasoners developed a sensitivity to conflict items. Hence, the first heuristic incorrect answer produced by system 1 was changed throughout the final answer which is the correct logical response produced by system 2. The default interventionist model predicts that the first heuristic response is incorrect and after additional consideration, incorrect responses might be corrected when the problems include a conflict item.

The mean values of reaction times within the direction of change categories were shown in the result section. It was found that reaction times increased in the initial incorrect and final correct responses category (see Table 10). According to the DI model, when the initial incorrect heuristic response produced by system 1 is changed to a correct response during the final response, the reaction time gets longer because system 2 thinking needs time to generate the correct response (Kahneman, 2011). The reaction times results of the "01" responses category in the experiment problems had the highest values for bat and ball problems (30.36) and syllogisms (19.17). This result is consistent with Bago & De Neys (2017, 2019) and Thompson et al., (2011). DI model prediction in this situation indicated that the incorrect response is corrected during the final response, and the reaction times increase because system 2 thinking is time consuming. Therefore, the result for "01" responses category reaction times in the experiment problems fit in the prediction of the DI model.

In the result section, a significant difference was observed between P1 confidence scale scores and P2 confidence scale scores. P2 confidence scale scores were significantly higher than P1 confidence scale scores. The mean values of confidence scale scores were indicated in the result section. It was revealed that when the initial response and final responses were matched, participants had high confidence (see Table 11). "Feeling of rightness" has a crucial role because when the initial response was not changed, the reasoners did not doubt their responses (Thompson et al., 2011). Therefore, the matching responses categories have higher confidence scale scores. The only one response category "10" responses within the bat and ball control problems did not fit in with this interpretation since just a trial was given as "10" responses.

Therefore, the mean value of confidence scale scores was seen as a “10” responses. These results are consistent with Bago & De Neys (2017).

5.1. Limitations

In this study, it was assumed that initial responses were given intuitively. However, “11” responses within the experiment problems’ direction of change category have still frequent. Therefore, the memory load and strict time limit might not be effective in the initial response stage for some participants. On the other hand, this procedure was effective but some participants may have had more difficulty. Therefore, they may respond randomly.

The memory load, which consisted of 4 digits, was used in this research. The 4 digits load could not be challenging for some participants. Therefore, their initial response may also have been deliberated. The digits in the memory load could be gradually increased when the participant responds correctly to the first problems. In addition, the 10 seconds time limit may not have been challenging enough, especially for participants with a high reading rate. The time limit could have been gradually reduced for participants who responded correctly in a short time.

5.2. Future Studies

The research (Ball et al., 2018; De Neys, 2006b; K. Stanovich, 2011; Keith E. Stanovich & West, 2000, 2008) revealed that individual differences differentiate the reasoning process. De Neys (2006b) indicated that the subjects with high span have better accuracy scores than the subjects with low span in the syllogisms which include conflict items. People, who have higher scores on general intelligence and working memory capacity tests, are shown better performance to corrected their first responses (De Neys, 2006b; Evans & Stanovich, 2013; Kahneman, 2011; K. E. Stanovich & West, 2003; Keith E. Stanovich & West, 2000). There is a positive correlation between fluid intelligence and working memory capacity (Conway et al., 2003). Individual differences have a crucial role in the reasoning process. In future studies, individual memory capacity and individual levels of attention parameter, which is an important component of working memory, can be measured. Then, reasoning questions can be asked to the participants through the two response paradigms. Thus, we can empirically test the effects of various mental activities on reasoning processes. In addition, the reasoning tasks which were used in this research can be applied to understand how the thinking processes in mental disorders such as attention hyperactive disorder, schizophrenia, and depression, where the affected individuals' cognitive processes such as attention and reasoning are heavily impaired. Comparisons can be made between the healthy population data that I obtained and clinical groups. Thus, it can be evaluated how thinking processes change in mental illnesses, and the dual process approach might be evaluated among clinical groups.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. *Psychology of Learning and Motivation - Advances in Research and Theory*, *2*(C), 89–195. https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. D. (1992). Working Memory Components of Working Memory Individual Differences in Working Memory The Slave Systems of Working Memory. *Science*, *255*(ii), 556–559. https://doi.org/10.4249/scholarpedia.3015

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109. https://doi.org/10.1016/j.cognition.2016.10.014

Bago, B., & De Neys, W. (2019). The Smart System 1: evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking and Reasoning*, *25*(3), 257–299. https://doi.org/10.1080/13546783.2018.1507949

Ball, L. J., Thompson, V. A., & Stupple, E. J. N. (2018). Conflict and Dual Process Theory. In *Dual Process Theory 2.0*. https://doi.org/10.4324/9781315204550-7

Baron, J. (2012). The point of normative models in judgment and decision making. *Frontiers in Psychology*, *3*(DEC). https://doi.org/10.3389/fpsyg.2012.00577

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. In *Trends in Cognitive Sciences* (Vol. 7, Issue 12). https://doi.org/10.1016/j.tics.2003.10.005

Cummins, R., & Fodor, J. (1985). The Modularity of Mind. *The Philosophical Review*, *94*(1), 101. https://doi.org/10.2307/2184717

De Neys, W. (2006a). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *Quarterly Journal of Experimental Psychology*, *59*(6), 1070–1100. https://doi.org/10.1080/02724980543000123

De Neys, W. (2006b). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*(5), 428–433. https://doi.org/10.1111/j.1467-9280.2006.01723.x

De Neys, W. (2017). Dual process theory 2.0. In *Dual Process Theory 2.0*. https://doi.org/10.4324/9781315204550

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3). https://doi.org/10.1016/j.cognition.2007.06.002

De Neys, W., & Pennycook, G. (2019). Logic, Fast and Slow: Advances in Dual-Process Theorizing. *Current Directions in Psychological Science*, *28*(5), 503–509. https://doi.org/10.1177/0963721419855658

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012

Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. In *Psychonomic Bulletin and Review* (Vol. 13, Issue 3). https://doi.org/10.3758/BF03193858

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. S. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking and Reasoning*, *11*(4). https://doi.org/10.1080/13546780542000005

Evans, J. S. B. T., Newstead, S. E., & Byrne, R. M. J. (2019). Human reasoning: The psychology of deduction. In *Human Reasoning: The Psychology of Deduction*. https://doi.org/10.4324/9781315785028

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Fodor, J. A. (2018). The Mind Doesn't Work That Way. In *The Mind Doesn't Work That Way*. https://doi.org/10.7551/mitpress/4627.001.0001

Frankish, K., & Evans, J. S. B. T. (2012). The duality of mind: An historical perspective. In *In Two Minds: Dual Processes and Beyond*. https://doi.org/10.1093/acprof:oso/9780199230167.003.0001

Franssens, S., & De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking and Reasoning*, *15*(2). https://doi.org/10.1080/13546780802711185

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2008). Cognitive Neuroscience: The Biology Of Mind (excerpt). In *Cognitive Neuroscience The Biology Of Mind*.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). Heuristics and Biases: The Psychology of Intuitive Judgment. In *Academy of Management Review* (Vol. 29, Issue 4).

Hastie, R., & Dawes, R. M. (2010). Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making. In *SAGE Publications* (Issue 2).

Kahneman, D, & Frederick, S. (2005). A model of heuristic judgment. *The Cambridge Handbook of Thinking and Reasoning*.

Kahneman, Daniel. (2002). Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice (Prize Lecture, December 8, 2002). *Les Prix Nobel: The Nobel Prizes 2002, December*, 449–489.

Kahneman, Daniel. (2011). Thinking fast and thinking slow. In *Farrar, Strauss and Giroux, New York, NY*.

Kahneman, Daniel, & Frederick, S. (2012). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. In *Heuristics and Biases*. https://doi.org/10.1017/cbo9780511808098.004

Kahneman, Daniel, & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3). https://doi.org/10.1016/0010-0285(72)90016-3

Kahneman, Daniel, & Tversky, A. (2013). The simulation heuristic. In *Judgment under Uncertainty*. https://doi.org/10.1017/cbo9780511809477.015

Keren, G., & Schul, Y. (2009). Two Is Not Always Better Than One: A Critical Evaluation of Two-System Theories. *Perspectives on Psychological Science*, *4*(6). https://doi.org/10.1111/j.1745-6924.2009.01164.x

Koehler, D. J., & Harvey, N. (2008). Blackwell Handbook of Judgment and Decision Making. In *Blackwell Handbook of Judgment and Decision Making*. https://doi.org/10.1002/9780470752937

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and Deliberate Judgments Are Based on Common Principles. *Psychological Review*, *118*(1). https://doi.org/10.1037/a0020762

La Pointe, L. B., & Engle, R. W. (1990). Simple and Complex Word Spans as Measures of Working Memory Capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6). https://doi.org/10.1037/0278-7393.16.6.1118

Manktelow, K., & Galbraith, N. (2012). Thinking and Reasoning. In *Thinking and Reasoning*. https://doi.org/10.4324/9780203115466

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, *17*(1). https://doi.org/10.3758/BF03199552

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*(4). https://doi.org/10.1037//0096-3445.130.4.621

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a Cognitive Theory of Emotions. *Cognition and Emotion*, *1*(1), 29–50. https://doi.org/10.1016/j.regpep.2012.06.013

Raoelison, M., & De Neys, W. (2019). Do we de-bias ourselves?: The impact of repeated presentation on the bat-and-ball problem. *Judgment and Decision Making*, *14*(2).

Raoelison, M. T. S., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*(December 2019). https://doi.org/10.1016/j.cognition.2020.104381

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1). https://doi.org/10.1037/0033-2909.119.1.3

Šrol, J., & De Neys, W. (2021). Predicting individual differences in conflict detection and bias susceptibility during reasoning. *Thinking and Reasoning*, *27*(1), 38–68. https://doi.org/10.1080/13546783.2019.1708793

Stanovich, K. (2011). Rationality and the Reflective Mind. In *Rationality and the Reflective Mind*. https://doi.org/10.1093/acprof:oso/9780195341140.001.0001

Stanovich, K. E. and W. R. F. (2006). Dual Processing in Reasoning. *Psychological Science, 17*(5), 428–433.

Stanovich, K. E., & West, R. F. (2003). "Individual differences in reasoning: Implications for the rationality debate? In *Behavioral and Brain Sciences* (Vol. 26, Issue 4). https://doi.org/10.1017/S0140525X03210116

Stanovich, K.E. (1999). Who is rational? Studies in individual differences in reasoning. In *Studies of individual differences in reasoning* (Vol. 20, Issue 3).

Stanovich, Keith E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate?-Open Peer Commentary-Understanding/acceptance and adaptation: Is the non- …. *Behavioral and Brain Science*, *23*(May 2019).

Stanovich, Keith E., & West, R. F. (2003). The rationality debate as a progressive research program. *Behavioral and Brain Sciences*, *26*(4). https://doi.org/10.1017/S0140525X03240115

Stanovich, Keith E., & West, R. F. (2008). On the Relative Independence of Thinking Biases and Cognitive Ability. *Journal of Personality and Social Psychology*, *94*(4). https://doi.org/10.1037/0022-3514.94.4.672

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3). https://doi.org/10.1207/s15327957pspr0803_1

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking and Reasoning*, *20*(2), 215–244. https://doi.org/10.1080/13546783.2013.869763

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3). https://doi.org/10.1016/j.cogpsych.2011.06.001

Toates, F. (2006). A model of the hierarchy of behaviour, cognition, and consciousness. *Consciousness and Cognition*, *15*(1). https://doi.org/10.1016/j.concog.2005.04.008

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2). https://doi.org/10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157). https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4). https://doi.org/10.1037/0033-295X.90.4.293

Wilson, T. D. (2022). 2 The Adaptive Unconscious. In *Strangers to Ourselves*. https://doi.org/10.4159/9780674045217-003