How to Cite:
Akbarifeizabadi, F., Nouri, R. K., & Mosmer, R. Get cognitively fatigued and be more confident! The effect of working memory capacity and cognitive load on meta-reasoning. *International Journal of Health Sciences*, 6191-6211. Retrieved from
https://sciencescholar.us/journal/index.php/ijhs/article/view/13460

Get cognitively fatigued and be more confident! The effect of working memory capacity and cognitive load on meta-reasoning

Faezeh Akbarifeizabadi
Shahid Beheshti University, cognitive science, Evin, Tehran, Iran
Email: f_akbarifeizabadi@sbu.ac.ir

Reza Kormi-Nouri
Örebro University, School of Law, Psychology and Social Work, Örebro, Sweden
Email: reza.kormi-nouri@oru.se

Reza Mosmer
Institute for Cognitive Science Studies, Philosophy of Mind, Pardis, Tehran, Iran
Email: Mosmer_r@icss.ac.ir

***Abstract*---**The goal of this study was to investigate the effect of cognitive limitations on intermediate judgments of confidence as one of the meta-reasoning components. This experimental study was conducted with 384 participants who were randomly assigned to two groups with different cognitive loads. The main task at the target level was the Tower of Hanoi, where the participants reported the level of their intermediate judgment of confidence every 45 seconds via a computerized questionnaire that popped up during the problem-solving process. The visual-spatial working memory capacity of individuals was measured with a forward and backward computerized Corsi block-tapping task. The participants in the high cognitive load group performed the dot matrix task before the Tower of Hanoi task. Two-way ANOVA found a link between the additional cognitive load before the main task and the participant's intermediate judgment of confidence during the problem-solving process. Visual-spatial working memory, on the other hand, did not have a significant effect on intermediate judgments of confidence, either alone or in conjunction with additional cognitive load.

Keywords*---intermediate judgment of confidence, working memory capacity, cognitive load, meta-reasoning.*

Manuscript submitted: 9 May 2022, Manuscript revised: 18 July 2022, Accepted for publication: 27 August 2022

Introduction

For centuries, the concept of metacognition has been taken into account by many researchers because of its widespread application in the fields of critical thinking, learning, and problem solving (Sodian, 2012). The "meta" in this term suggests that both the subject and object of mental processes is cognition. In order to investigate the link between these two phenomena, cognition and metacognition, consider subjects who become trapped in a maze and attempt to escape. The cognitive system immediately begins processing the visual and special data. However, as Chalmers (Chalmers, 1995:212) argued, what took place within their minds that resulted in this novel experience, and how can we explain or describe their mental representation? The researchers used meta-cognition assessment to determine the answer to the preceding question by asking participants to report the level of their confidence. For example, in the maze scenario, subjects were asked to assess their level of confidence in the path they chose. As a consequence, by comparing their metacognitive judgement accuracy to what has been observed, it is possible to compare their internal representation of their own cognitive system activity with the real performance of their cognitive system (Grimaldi, Lau, and Basso 2015; Pouget, Drugowitsch, and Kepecs 2016).

However, it appears as though the interaction between metacognition and cognition is bidirectional. At first metacognitive abilities do not develop until the cognitive system does; after that, it is thought that having strong metacognitive abilities improves cognitive performance (Akturk and Sahin, 2011; Lai 2011). In the subject of rationality and, more particularly, logical problem-solving, metacognition may be characterized as possessing relevant information, recognizing the time required, and using the most suitable method to solve or learn issues, among other things (Metcalfe and Shimamura, 1994). According to various systematic reviews that examined the relationship between metacognition and academic success, such as arithmetic proficiency, meta-reasoning, an important component of metacognition, may also play a role in learning, thinking, and problem-solving (Muncer et al., 2021).

The concept of meta-reasoning, which was defined as a subset of metacognition by Ackerman and Thompson (2017) in accordance with Nelson and Naren's meta-memory model (1994). Meta-reasoning is when a person observes their entire process of self-performance in logical problem-solving or reasoning on a cognitive level. At the meta-level, monitoring entails evaluating how individuals are confident about what they are currently doing at the cognitive level. One of the meta-reasoning components during the problem-solving process that these authors proposed is an intermediate judgment of confidence (IJC).

IJC is defined as self-comparison of the subject's current confidence with his or her ideal level of confidence, which means continuously assessing the subjects' confidence in their own performance on cognitive grounds that may prompt them to control their cognitive activity during logical reasoning or problem-solving (Ackerman and Thompson 2017). As previously stated, IJC integrates two cognitive dimensions as a meta-reasoning component. The subjective part is aligned with the meta-memory model at meta-level, while the objective one refers to rationality at cognitive level.

Thinking as a high-level cognitive process includes three main components: reasoning, problem solving, and decision-making. According to the definition of meta-reasoning, components of rationality such as reasoning and judgment follow the dual processing model. In this model and all versions of this theory, there is one commonality that refers to the idea of two dissimilar processes involved in thinking. In the present study, we called these two processes "system 1" and "system 2" by referring to Kahneman's work (2003). The first system is fast and intutive, while the second system is deliberative and slow because it is limited by cognitive resources such as working memory capacity (Stanovich, 1999). System 1 is based on mental shortcuts and draws on experience, while system 2 is based on cognitive abilities (Evans, 2003). According to the findings of a prior research study conducted by Hambrick and Engle (2003), working memory plays an essential role in the process of reasoning and problem-solving. The main goal of this research is to explore, either directly or indirectly, the link between meta-reasoning as a metacognitive component and working memory as a cognitive component in the human mind.

Alan Baddeley (1974) defined working memory (WM) as a multicomponent system, with the visual-spatial sketchpad subsystem storing visual data and information about color, shape, spatial placement, and motion and supplying it to the central operating system (Turnbull, 2012). Working memory is a fundamental memory element for cognitive functions; it is an active memory system that momentarily retains information and is integrated with higher-level cognition such as long-term memory, language, reasoning and problem solving (Baddeley, 1974, 2012). As Cowan (1995) demonstrated, working memory capacity and focused attention are both located in the parietal lobes, and the findings of the ERP study revealed that the presence of visual cues in mathematical problem solving, such as the Raven Progressive Matrices, can aid a person in diverting their attention away from the disruptive factors that cause them to forget a mathematical problem (Cowan, 1995; Williams, and Woodman, 2012; Wiley, and Jarosz, 2012).

Barrett and colleagues (2004) argue that the capacity of working memory is allocated to the keeping and management of information in reasoning tasks until interruption is controlled, and that system 2 is responsible for information processing throughout these activities. In other words, the limited source of attention in the working memory model, which is contained in its core central operating system, is used to manage and filter information during control processes, especially when distraction occurs and focused attention is needed. Therefore, individuals' working memory capacity differences are a consequence of the engagement of system 2 (slow) in the operation of centeral system (Barrett, Tugade, and Engle, 2004). However, cognitive load is another approach that has been used in the literature to investigate the impact of working memory capacity on the learning process, regardless of individual differences (Kirschner, Paas, and Kirschner, 2009).

Sweller (1988) has also demonstrated that working memory plays an important role in learning while solving mathematics problems, either directly or indirectly through cognitive load, which speeds up and facilitates information processing in working memory by employing various types of cognitive load, such as 1) intrinsic load, defined as the number of elements that must be concurrently processed in working memory for schema construction; 2) extraneous load, caused by instructional techniques that require learners to engage in working memory activities; and 3) germane load, caused by beneficial cognitive processes (Jaeger, 2017; Paas, 2003; Sweller, 2011). A combination of cognitive load and time pressure was used by the researchers to activate system 1 and suppress system 2, in line with the link between cognitive load and working memory capacity (Bago and De Neys, 2020).

It is also assumed that all types of feelings and judgments that are components of the meta-reasoning monitoring process act fast, intuitively, and based on system 1. Except in cases where the IJC is a repetitive and recursive action over time at the meta-level, the subject needs to use his or her working memory for retrieving the ideal confidence level and comparing it with the current one, while working memory capacity is limited and may be at the service of system 2 for reasoning tasks at a cognitive level (Feeney and Thompson, 2015). By addressing this situation, the following issue arises: Is system 1's speed for IJC slower than its speed for other meta-reasoning components? How could the speed difference between systems 1 and 2 be measured? Is there a speed spectrum in which these two systems are placed at the maximum and minimum points? How do these two systems interact with each other?

Several neurological studies have demonstrated that system 2, unlike system 1, is unique to humans and is involved in conceptual thinking (Feeney and Thomson, 2015). There are some aspects of the dual process theory that are unclear and have been criticized for many years, such as the manner in which systems 1 and 2 interact with each other and how these two systems are distinct from each other (Evans and Stanovich, 2013). One of the well-known criticisms of Kahneman's dual processing theory is the constant nature of communication between the two systems. For example, it is claimed that only system 1 processes at the metacognitive level, but it is unclear how this activity is initiated. Van Rooij and colleagues (2016) suggest that the relationships between systems 1 and 2 should be dynamic, not static. This means that they sometimes work together and in parallel, but most of the time, they work sequentially, with cognitive load acting as a barrier between these two phases (Van Rooij and Favela, 2016).

In the present study, we wanted to address these critiques with the following research questions:

Does IJC only use system 1 (intuitive processing), or does comparing current and ideal confidence need working memory capacity and is based on system 2 (deliberate processing)?

Could increasing the cognitive load have an impact on the participants' IJC at the metacognitive level?

In order to escape the limitations produced by critiques of the communication between systems 1 and 2 in the dual processing paradigm, we assume in this research that the interaction between systems 1 and 2 is parallel and continuous. As a result, the IJC's functioning is based on system 1 at the metacognitive level, while system 2 is involved in the problem-solving process at the cognitive level.

With this information in mind, we developed the following hypothesis regarding how two distinct level of visual-spatial working memory capacity (low vs. high) would have an impact on IJC (with or without the precent of additional cognitive load) during the problem-solving process.

Firstly, we assume the average IJC of the participants would not differ significantly between the two groups (with and without additional cognitive load). Then participants with distinct levels of visual-spatial working memory capacity would all get the same average IJC score, regardless of their level of working memory. In the end, it makes no difference whether you have a high or low visual-spatial working memory when it comes to the effect of cognitive load on the IJC.

To test these hypothesis, and because of the repetition of IJC throughout the problem-solving process, as well as the influence of time pressure with cognitive load on system 1 activation, we used the Tower of Hanoi (TOH) as our main task without any time limit and without offering any reward or punishment for the following reasons: 1. it is a sequential puzzle, and the condition of the issue at each iteration is the consequence of the preceding thought process, which demands a phase-by-phase approach over time; 2. Our goal was to diminish the impact of language on judgment and to eliminate any biases that could arise as a consequence of using a visual task for problem solving; and 3. Because this cognitive activity has been empirically shown in many kinds of puzzles and games currently on the market, the TOH is considered to be a good option for controlling the effects of familiarity.

Furthermore, individuals with low working memory capacity perform slowly in counting tasks and quickly in complex tasks, according to the findings of a study conducted by Souza and Oberauer (2015) to compare the forgetting process in tasks related to visual working memory under absolute and relative time conditions. By creating a computer-based version of the TOH problem with no completion deadline, we try to retain control over the process and ensure that everyone, regardless of their visual-spatial working memory capability, experiences the identical circumstances. This software frequently collects the participant's IJC throughout the problem-solving process at a specified period of time, as inspired by the Metcalfe and Shimamura (1994) technique for measuring metacognition components and taking into consideration the stable parallel foundation of systems 1 and 2 interaction.

Additionally, by designing this software, we wanted to lay a foundation for future research on one of the hottest areas in cognitive neuroscience: the relationship between metacognition and functional cognitive impairments (Larner, 2021).

Previous research found that a rapid change in emotion from pleasant to unpleasant (happy to sad) and vice versa impairs the function of the meta-reasoning system for 5 minutes (Folwarczny, Kaczmarek, Doliński, and Szczepanowski, 2018). In addition, there is more evidence in the articles about how an individual's emotional state, especially their level of arousal, can change their judgment from positive to negative or neutral (Manstead, Wagner, and MacDonald, 1983; Clark, Milberg, and Erber, 1984). In this respect, we decided to measure the emotional state of the participants before and after the IJC evaluation in order to control the influence of emotions.

Methods

Participants:

This study was conducted with 384 participants, including 233 females and 151 males. The participants included 95 postgraduates, 21 graduate students and 268 undergraduates; the participants were randomly assigned into two groups, one without a significant level of cognitive load (labeled as Group 1) and the other with an additional cognitive load (labeled as Group 2). Studies have shown that metacognitive skills are among the abilities that grow slowly and require the development of executive function, verbal abilities, and inhibitory control (Lai, 2011). Therefore, to control for aging effects, we had two inclusion criteria: the participants were between 20 and 40 years old and had no history of chronic or acute diseases. The participants who believed that our reasoning task (i.e., the TOH problem) could not be solved and failed to solve the problem in less than 45 seconds were excluded and replaced by new participants. The participants were randomly assigned numbers from 0 to 480; then, the participants were selected from a numerical list, and they alternatively joined one of the two groups. Therefore, everyone had an equal chance of participating in each group, and the group allocation in this experiment was done completely randomly.

Materials:

1- Tower of Hanoi (TOH):

Previous research in Thomson and Ackerman's (2017) meta-reasoning category used logical tasks (such as Raven matrices, remote association test, and water jar problem), which can be done rapidly, or at the very least in a short amount of time. However, we used the "Tower of Hanoi" (TOH), which naturally takes time to resolve, since the actions and tactics used by participants while solving the problem result in the formation of new obstacles.

The TOH is a mathematical puzzle with three rods and n discs. The objective of the puzzle is to move the entire stack to another rod by obeying the following simple rules: only one disc can be moved at a time, and two discs may not be placed on top of a smaller disc. Thus, since the TOH has a starting point and a goal to reach, it is a goal-directed task. It is also a problem-solving task since there are obstacles that become apparent when approaching the goal by applying the above rules. To solve this problem, various situations arise for which different solutions have been offered. In this study, the TOH puzzle is considered the main task to solve the problem at the cognitive level.

To examine the IJC, the participants were asked to judge their confidence in their performance on a 5-point scale every 45 seconds during the problem-solving process (Figure 1). The questionnaire disappeared from the monitor after 15 seconds, and the participants could not move any discs at the target level unless they chose one answer and clicked the continue button.

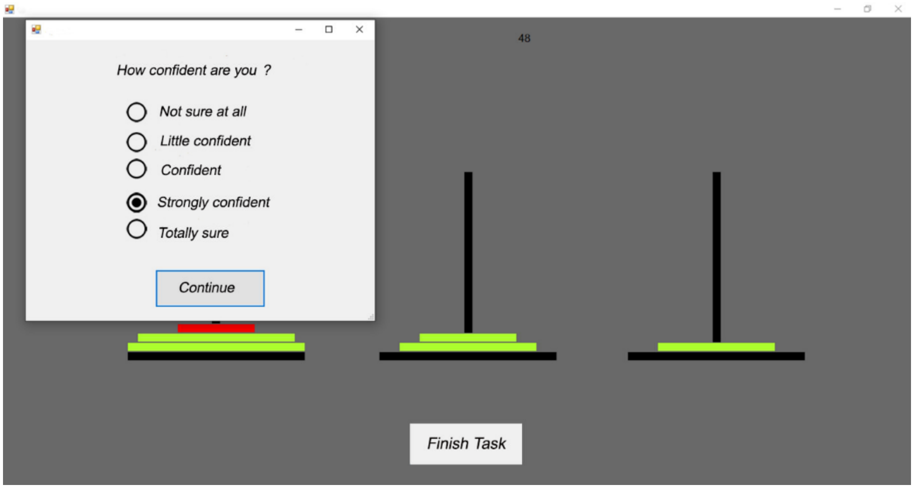

Fig. 1. Tower of Hanoi at the cognitive level with the IJC questionnaire at the meta-level

2- Corsi Visual-Spatial Working Memory Test

In Iran, a Corsi block-tapping task was normalized as a subtest in the Wechsler memory package for the Iranian population between 16 and 40 years old (Saed et al., 2008). The validity of the Iranian version of Corsi's task was 79%, and its reliability coefficient for working memory was 80%. To measure visual-spatial working memory capacity, Corsi visual-spatial working memory software made by the SINA Research Institute of Cognitive Behavioral Sciences and Research (Khodadadi and Amani, 2018) was used. In the computerized Corsi test, several blocks appear on the screen, and some of these blocks are turned on and off in turn. The subjects were asked to memorize the location of the blocks that had been turned on and then click on the blocks in the same order.

The Corsi block-tapping test involves mimicking a researcher as they tap a sequence of up to nine identical spatially separated blocks. The initial sequences are simple, usually using two blocks, but become more complex until the participant's performance degrades. This number is known as the Corsi span, and the average is approximately 5-6 for the typical "normal" human participant. If any participant completed both sequences, he or she could go to the next sequence, and if he or she made a mistake in either sequence, the test would stop. In this test, memory was first measured in forward order, and subsequently, the participants continued in the backward order, in which they should click on the lit blocks with the mouse according to the photo order.

For the scoring of the Corsi visual-spatial working memory test, the total scores of the forward and backward orders were added, and the total score for each order was obtained by multiplying the number of correct attempts by the number of test phases.

3- Self-Assessment Manikin (SAM) Questionnaire

The three-dimensional model of the Self-Assessment Manikin (SAM) scale, developed by Bradley and Lang in 1980, is a pictorial rating system to obtain self-assessments of experienced emotions on the dimensions of affective pleasure, dominance, and arousal. A paper-pencil or computer-based version can be used in an experimental procedure. In Iran, the validity and reliability of the arousal and pleasure domains of this questionnaire have been reviewed and confirmed (Nazari et al., 2011).

In this study, immediately before and after the TOH task, the subjects were presented with the self-assessment questions, and they were asked to express their current emotions by choosing one picture among the pictures ranging from happy (1) to sad (5). In addition, they were asked to score their level of arousal participants. The emotions ranged from ecstasy to drowsiness or boredom (1–5).

4- Dot Matrix Task

The dot-matrix task was modeled after Miyake et al. (2001). A matrix equation was verified, and then a dot location displayed in a five-by-five grid had to be remembered. The matrix equation was presented for a maximum of 1000 milliseconds. Following that, a list of mixed letters E and F was shown to the participants, and they were instructed to find letter F among the others. Then, a five-by-five grid of empty spaces appeared on the monitor, and the participants were instructed to recall the memorized dots by clicking on the relevant locations with the mouse.

The reliability and validity of the computer version of this task have been investigated by Sliwinski et al. (2018). To confirm the reliability of this task, the intergroup variance was obtained at 0.77 and the intragroup variance at 1.19 in our data, which showed the high reliability of the model. In terms of validity, it has been shown that dot memory tasks, like other working memory tasks, are correlated with age.

5- PAAS Standard Cognitive Load Assessment Questionnaire

The PAAS questionnaire has 4 questions scored on a 9-point Likert scale (from easy to too difficult, 1–9) (Paas, 1992). The total score could vary from 4 (minimum cognitive load) to 36 (maximum cognitive load). The reliability coefficient of this questionnaire, based on Cronbach's alpha coefficient, was 0.85. The validity of the questionnaire has been confirmed by Iranian professors and experts in this field (Ahadi and Soleimani, 2014).

6- Stroop Color Task

The Stroop test from the MEHRSA Publications' PCAP cognitive assessment package (Rezapour et al., 2013) was used in this study. First, in the baseline condition, blue, yellow, red, and green colored blocks were shown on the screen, and the participants were asked to say the names of these colors in the list from the top right column to the bottom left column aloud. The number of incorrect responses and the reaction times of the participants were recorded by the experimenter. Second, in the control condition, a list containing the written form of color names (red, blue, yellow, and green) that were presented in the same color was displayed. The participants were instructed to read the written words in the same manner as in the experimental stage. Third, in the experimental condition, similar to the control condition, the written forms of the color names were shown. However, the color names were presented in a different color. For example, the word "red" was written in green ink. The participants were instructed to name the text's color aloud in the same order as before.

The Stroop test was scored using two outputs per subject: 1-the total of the errors in the second and third stages; 2-the interference score is derived from the difference in reaction time between the second and third stages.

Procedure:

For Group 1 (with low cognitive load), in the first phase, the participants were tested with the Corsi block-tapping test (with both forward and backward tasks), which is a psychological test for assessing visual-spatial working memory capacity. In the second phase, the Self-Assessment Manikin (SAM) five-point scale was used to control for the participants' emotions, and their responses were recorded. In the third phase, the instructions and implementation for the Towers-of-Hanoi (TOH) task were explained to the subjects both orally and in writing. After ensuring that the participants correctly understood the task instructions, the process of solving the TOH problem began for the participants. In the fourth phase, the Self-Assessment Manikin (SAM) scale was administered to reassess the emotions of the participants after the task was performed. They were then asked to give a brief explanation of the method and strategy they used to solve the problem, to state whether they had prior knowledge about the TOH problem or a similar version such as "The Tower of London Test" (TOL), and to report to us how much they forgot about whatever they were supposed to do with regard to the IJC questionnaire during the task.

For Group 2 (with high cognitive load), the data collection procedure was similar to that used in Group 1 with only two differences: first, in the second phase, they were asked to perform an additional task, namely, the "dot matrix cognitive training" task, to increase their cognitive load. The second difference was related to the TOH task. Although the participants in Group 2 received the same instructions and implementation as Group 1 in the fourth phase, the use of the middle bar was not mentioned in the instructions. Thus, the structure of the instructions used with the second group was more complex than for the first group. The remaining parts of the procedure (e.g., using the Self-Assessment Manikin scale before and after the Tower of Hanoi task and the questionnaire about their prior knowledge) were administered in a similar manner as in Group 1. The procedures used in Group 1 and Group 2 are depicted in Figure 2. To address the assessment of meta-reasoning, the TOH task was delivered through a computer and software. We developed custom software that required the participants to input information on their IJC at certain points throughout the TOH problem-solving process. At minimum, the components of time, attention, concentration disturbances, environment conditions, and cognitive level of the test were the same across all participants, and some attempts were made to eliminate intervening factors. Computer-assisted data collection resulted in a decrease in processing speeds and human error rates. Prior to the main study, a pilot study was conducted with 24 volunteer participants (all university students) ranging in age from 20 to 40 years old. It was demonstrated that performing the spatial N-back test before and after the main TOH task could result in bias and errors. Thus, we chose the Corsi block-tapping test to evaluate visual-spatial working memory capacity.

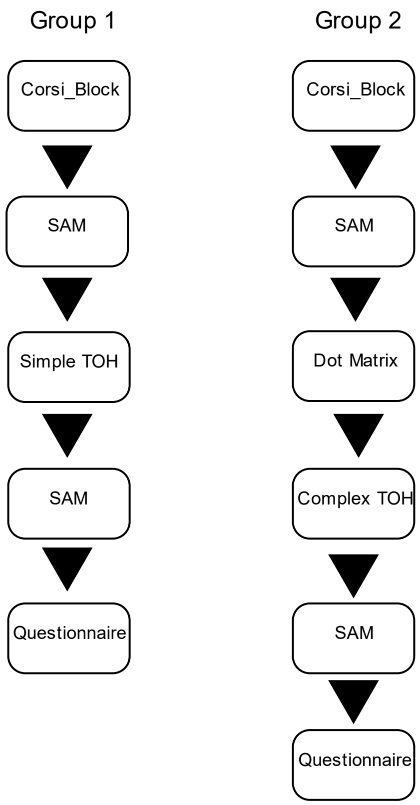

Fig. 2. Experimental pathway for the high and low cognitive load groups

According to Gwizdka (2010), performance-based approaches could be used to assess the cognitive load on a user's primary task by using Stroop reaction time as a secondary task. Therefore, in the pilot, the PAAS Standard Cognitive Load Assessment Questionnaire (1994), as well as the Stroop test, were used to compare and ensure cognitive load differences between the two groups of participants.

Following the pilot study, the three questions listed below were added as the final phase for both groups based on participant comments regarding the influence of the IJC questionnaire, which appeared on the screen every 45 seconds, and on forgetting the strategy they had considered. The purpose of the second and third questions was to control for the effects of familiarity and chunking. The participants may have recognized the TOH or TOL without knowing their names, and they may have known how to efficiently transfer sub towers from a particular source peg to a destination peg in a single phase, known as chunks (Ruiz and Newell, 1989).

1. In your opinion, when you answered the questions related to the confidence percentage of your performance, did it make you forget what you wanted to do? 1. not at all; 2. little; 3. moderate; 4. a lot

2. Were you familiar with the problem-solving task (TOH) used in this study or a similar task (TOL)? 1. none; 2. TOL; 3. TOH; 4. both.

3. Please explain the strategy that you used to solve the Towers-of-Hanoi problem.
We planned to use skin conductance sensors as a physiological technique in the pilot study to assess cognitive load and participant stress level, but due to the pandemic situation and considering the participant's health, we decided to omit this feature of the experiment.

Results

Data preprocessing:

To evaluate the difference in cognitive load between the two groups, we conducted a pilot study with 24 participants (12 participants in each group). This was established via the use of the PAAS and Stroop tasks.The mean score of the PAAS questionnaire for Group 1 was 13.58, which was lower than the mean score of the Pass questionnaire of Group 2. In addition, the mean of Stroop task interference scores (65.08 for Group 1 and 100.58 for Group 2) reveal that Group 2 participants faced an additional cognitive load compared to Group 1 participants.
To reach a result with a confidence of 95% and an accuracy of at least 5% (plus or minus), we calculated the sample size using Cochran's method with the data from the pilot research. Thus, a random sample of 384 students from our target population should be sufficient to provide the degree of confidence required for the original research.

Following the pilot study, we examined how both groups of persons performed in the TOH problem-solving process to see if the additional cognitive load had an effect on their performance. There were 142 participants who correctly solved the TOH and 242 participants who could not. The t test findings in Table 1 indicate how long each group member took to complete the TOH, how many errors were made, and how many disc movements were performed. The outcomes demonstrate that participants in Group 1, who had less cognitive load, did better than those in Group 2.

Table 1. Total TOH time, number of errors and number of disc movements for Group 1 and 2 participants.

Group Statistics					Levene's Test for Equality of Variances		t-test for Equality of Means		
	Group	Mean	Std. Deviation	Std. Error Mean	F	Sig	t	df	Sig. (2-tailed)
number of errors	1	4.29	5.78	0.41	21.98	0.00	-3.22	382	0.00
	2	6.70	8.58	0.61					
number of disc movements	1	210.34	254.10	18.33	3.56	0.06	-1.64	382	0.1
	2	253.43	258.22	18.63					
Total TOH time	1	625.34	714.46	51.56	4.23	0.40	-1.82	382	0.06
	2	763.92	773.10	55.79					

WM capacity was measured using the forward and backward order tasks. The Kolmogorov–Smirnov test was used to examine the normality of the data distribution on the WM capacity of 384 subjects (without any missing data). Since the significance value was0.04 which was lower than 0.05, the data did not show a normal distribution. Then, a receiver operating characteristic (ROC) curve was drawn to evaluate the sensitivity and specificity of selected cutoff points that resulted in no false-positivesc and was a good choice. This curve compared the cutoff point of the Corsi visual-spatial WM capacity scores with the previous spatial N-back task scores of 24 participants in the pilot study. Eleven of the subjects with a cutoff point of 74 were identified as having high WM capacity, and 13 of them were identified as having low WM capacity.

According to the cutoff point, the subjects with a visual-spatial WM score above 75 were categorized into the group with high WM capacity. The participants with a total score of WM capacity of 74 or lower were categorized into the group with low WM capacity.

The questionnaire was presented every 45 seconds, so the number of judgments varied across participants depending on the time they used to solve the problem. Therefore, it was difficult to compare the total IJC of the participants within a group or between the two different groups. In addition, some subjects left some questions related to their intermediate confidence during the problem-solving process unanswered, and they were labeled "Null" in the answering system. The null label indicated that the subject paused for 15 seconds and failed to answer the question (i.e., it was an unanswered question). Therefore, the more unanswered questions, the longer it took to solve the problem. For this study, it was assumed that these missing data can provide us with information about the subjects' behavior at the meta-level or target level and its relationship with their performance. The amount of null data was calculated and stored for each subject rather than eliminating or replacing it with a mean score.

To solve the first problem mentioned above, the preprocessing operation for all subjects was individually performed for the answers to the question regarding intermediate confidence of their performance The IJC options were scored as follows: 0 for not sure at all, 25 for little confidence, 50 for confidence, 75 for strongly confident, and 100 for totally sure. Then, for each participant, a score between 0 and 100 was assigned based on the most frequently occurring response (IJC options) to the IJC question.

3.2. Analyzing the Hypothesis

Two-way ANOVA was performed to analyze the effect of the additional cognitive load and visual-spatial WM capacity on intermediate judgment of confidence during the TOH problem-solving process. The number of participants with various visual-spatial WM capacities is shown in Table 2 as two groups: those who experienced no extra cognitive load (Group 1) and those who encountered additional cognitive load (Group 2).

Table 2. Mean IJC scores for participants with high and low visual-spatial WM capacity was distributed between groups one and two.

Group	WM capacity	Mean	Std. Deviation	N
1	Low	39.65	34.07	99
	High	34.14	32.71	93
	Total	36.98	33.44	192
2	Low	53.57	33.15	105
	High	57.76	31.96	87
	Total	55.47	32.60	192
Total	Low	46.81	34.23	204

	High	45.56	34.36	180
	Total	46.22	34.25	384

The results of Levene's test in shows that the variances for IJC in participants with low and high visual-spatial WM capacity in both groups were equal (F=0.04, P=0.98).

As illustrated in Table 3, two-way ANOVA revealed that the two cognitive load groups were significantly different ($P < .001$), indicating that the high cognitive load group had worse performance than the low cognitive load group. There was no main effect for the WM capacity variable, indicating no difference between participants with low and high WM capacity ($F = 0.03$, $p = 0.84$). According to a simple main effects analysis, the interaction between visual-spatial WM capacity and cognitive load was not statistically significant ($p = 0.15$).

By adding the additional cognitive load to both individuals with high and low visual-spatial WM capacity during the TOH problem-solving process, IJC increased (see Figure 3).

Table 3. Results of two-way ANOVA with IJC as the dependent variable and groups and WM capacity as the independent factors

Source	Sum of Square	df	Mean Square	F	Sig.
Group	0.01	1	33663.18	30.86	0.00
WM capacity	46.24	1	41.58	0.03	0.84
Interaction	1.79	1	2244.26	2.05	0.15

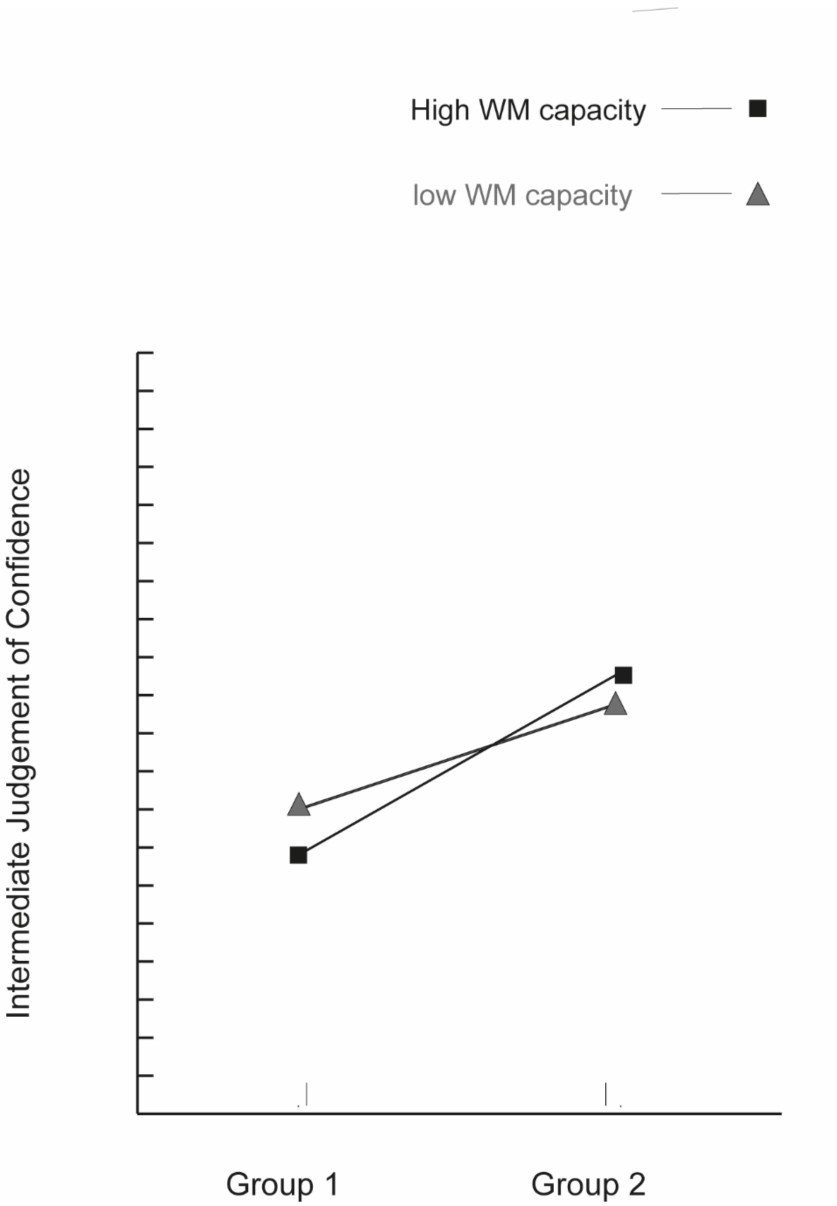

Fig. 3. Effects of cognitive load (Group 1 with low cognitive load and Group 2 with high cognitive load) and WM capacity on the IJC during the TOH problem-solving process

To control for the effect of emotions in this study, the level of pleasure and arousal of the participants was measured by the Self-Assessment Manikin (SAM) scales before and after the TOH task: the differences between the scores before and after each test were compared. The extent of pleasure and arousal changes of the participants was examined as described in Table 4. The skewness and kurtosis of the data related to the difference between the scores of pleasure and arousal was less than 3, which indicated that these data were normally distributed. By considering their mean values, which were approximately zero for both groups, it can be concluded that the levels of pleasure and arousal of subjects during the TOH problem task were not changed.

Table 4. Descriptive statistics of the difference in Self-Assessment Manikin (SAM) scores for each subject from before to after the task

	Frequency	Mean score difference	Standard deviation	Skewness	Kurtosis
Pleasure	384	0.36	1.10	0.13	1.89
Arousal	384	0.03	0.86	0.09	2.16

In this study, the concept of emotion was assessed with only two components, pleasure and arousal, and the component of dominance in the Self-Assessment Manikin (SAM) scale was not evaluated due to the sensitivity of judgment and bias. The evaluation of pleasure and arousal did not significantly change from before to after the TOH task, but in a small number of participants (36), emotional behavior was observed as cautious behavior, indicating more inhibitory control over their actions, and impulsive behavior, indicating less inhibitory control over their actions. Finally, by observing the rate of forgetfulness reported by participants at the task's conclusion, it is clear that the majority of participants with lower confidence reported having forgotten the strategy they were using during the TOH problem-solving process as a result of the IJC question popping up every 45 seconds and distracting them (Table 5).

Table 5. Frequency of participants' reported levels of forgetfulness depends on their IJC

		Forgetting the strategy				Total
		1-not at all	2- little	3-moderate	4-a lot	
IJC	Not sure at all	8	11	22	38	79
	Little confidence	10	23	41	21	95
	Confident	18	28	23	15	84
	Strongly confident	15	23	19	6	63
	Totally sure	17	28	15	3	63
Total		68	113	120	83	384

Discussion

This study aimed to investigate the effect of working memory capacity and cognitive load on intermediate judgments of confidence as a component of meta-reasoning. The findings based on the TOH problem-solving data indicated that Group 1 participants made fewer responses in a shorter time period and committed fewer mistakes and, therefore, performed better in the problem-solving process than the participants in Group 2 who were expected to experience additional cognitive load.

Two-way ANOVA rejected our null hypothesis about the average IJC being equal for participants in both groups with and without additional cognitive load during the TOH problem-solving process, and the results of this study showed that the average IJC for participants who experienced additional cognitive load increased. Additionally, having a poor or high capacity for visual spatial working memory had no direct or indirect influence on participants' IJC.

In light of this, the results of this study showed only how the presence of cognitive load affected participant's IJC about their cognitive performance at the metacognitive level and the speed with which they solved problems at the cognitive level. However, to more precisely investigate the relationship between cognitive load and the thinking speed in systems 1 and 2 in a dual processing model, a scale for measuring thinking component speed at both the cognitive and metacognitive levels was needed.

For this reason, and based on what De Neys (2006) mentioned regarding changing the speed of the reasoning process when a subject faces a contrast in context, we propose two new parameters based on each participant's collected data for future study. The first parameter is called the speed of intermediate judgment, and the second parameter is the IJC change rate. In this way, we can quantify the act of IJC for each participant during the problem-solving process based on how fast or slow the response is and based on a better measure of its fluctuation in time.

IJC change rate=number of times IJC changed × number of times it takes more than 2 seconds to answer IJC question

$$IJC\ speed = \left(\frac{\Sigma(IJC's\ score\ - next\ IJC's\ score)}{total\ time\ to\ solve\ TOH(sec)}\right)$$

Most individuals with the highest rate of strategy forgetting throughout the problem-solving process had the lowest level of IJC ("Not sure at all"; score=0), and vice versa. The act of reporting a forgetting strategy is a kind of observation of memory function by the individual and thus is expected to belong to the meta-level and to be a metacognitive element, whereas the act of reporting a forgotten strategy takes place after the forgotten occurrence. This report also mentioned the individual's perception of their memory's inability to retrieve information or a dysfunction of working memory, in which case, this recall of forgetfulness itself belongs to the meta-level because it is an action whose object is the action itself. Given that meta-reasoning is defined in the meta-memory framework, the significant correlation between a report of forgetting the strategy and IJC suggests a direct link between metacognitive elements at the meta-level. Therefore, regarding the act of the monitoring process at the memory level, the participants expressed a level of forgetfulness during the problem-solving process, and it is suggested that the relationship between meta-memory and meta-reasoning be examined in future studies.

In addition to its vital function in movement planning, visuospatial working memory capacity is limited by its complexity; hence, our visual working memory assessment may have been insufficient to discriminate between high and low WM competence (Kemps, 2001).

Limitations:

The conceptual ambiguity in the dual processing model and how systems 1 and 2 interact with the components of working memory theory was one of the challenges that limited us to employing visual tasks in this study. While we recognize that meta-reasoning has only been defined for a short time, there is still a lack of evidence and studies on the subject. While researching this topic, we were confronted with several issues and ambiguities.

Metacognition research is subjective and internal, and there is no accurate assessment tool for examining metacognitive components other than observation and self-expression (Lai, 2011), which is why we used a questionnaire.

We could not employ physiological methods to adequately assess a more accurate measure because of the experiment's scheduling conflict with pandemic conditions.

The relationship between working memory capacity and higher-level cognition is mostly unclear (Barrouillet, Lépine, & Camos, 2008).

Conclusion:

The results of the present study showed that for problems that require stages and need time to be solved, with increasing cognitive load at the target level for all participants, the level of IJC increased, and participants became more confident in their performance. However, differences in the visual-spatial working memory capacity of individuals did not directly affect their IJC. Based on the observations and collected data during this study, two parameters called IJC speed and IJC change rate were proposed.

For future studies, assessing correlations of IJC speed and IJC change rate with intermediate judgment confidence and additional cognitive load was suggested. Furthermore, we need to investigate whether it is possible to predict future problem-solving situations based on the pattern of changing IJC of participants at the meta-level.

Statements and Declarations

The Institute for Cognitive Sciences Studies (ICSS) granted permission to conduct this study. This was done as part of a Shahid Beheshti University research project.

The authors declare that they have no conflicts of interest or personal relationships that could have influenced the research presented in this paper.

All procedures used in studies involving human participants complied with the institutional and/or national research committee's ethical standards, as well as the 1964 Helsinki Declaration and its subsequent amendments or comparable ethical standards.

Informed consent was obtained from all individual participants involved in the study.

The datasets generated and analysed during the current study are not publicly available due the fact that they constitute an excerpt of research in progress but are available from the corresponding author on reasonable request.

References

Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. Trends in cognitive sciences, 21(8), 607-617.

Ahadi, F., & Soleimani, M. (2014). Comparing the effect of two teaching methods by presenting a completely solved educational example. Iranian Journal of Medical Education, 14(4), 291-302.

Akturk, A. O., & Sahin, I. (2011). Literature review on metacognition and its measurement. Procedia-Social and Behavioral Sciences, 15, 3731-3736.

Baddeley, A. (2012). Working memory: theories, models, and controversies. Annual review of psychology, 63, 1-29.

Bago, B., & De Neys, W. (2020). Advancing the specification of dual process models of higher cognition: A critical test of the hybrid model view. Thinking & Reasoning, 26(1), 1-30.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. Psychological bulletin, 130(4), 553.

Barrouillet, P., Lépine, R., & Camos, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes?. Psychonomic Bulletin & Review, 15(3), 528-534.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. Journal of consciousness studies, 2(3), 200-219.

Clark, M. S., Milberg, S., & Erber, R. (1984). Effects of arousal on judgments of others' emotions. Journal of Personality and Social Psychology, 46(3), 551.

Cowan, N. (1995). Attention and memory: An integrated framework New York: Oxford University Press.

De Neys, W., (2006). Dual processing in reasoning: Two systems but one reasoner. Psychological science, 17(5), 428-433.

Evans, J. S. B. (2003). In two minds: dual-process accounts of reasoning. Trends in cognitive sciences, 7(10), 454-459.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. Perspectives on psychological science, 8(3), 223-241.

Feeney, A., & Thompson, V. A. (Eds.). (2015). Reasoning as memory. Psychology Press.

Folwarczny, M., Kaczmarek, M. C., Doliński, D., & Szczepanowski, R. (2018). Emotional see-saw affects rationality of decision-making: Evidence for metacognitive impairments. Acta psychologica, 186, 126-132.

Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. Neuroscience & Biobehavioral Reviews, 55, 88-97.
Gwizdka, J. (2010, August). Using stroop task to assess cognitive load. In Proceedings of the 28th Annual European Conference on Cognitive Ergonomics (pp. 219-222).
Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. The psychology of problem solving, 176-206.
Jaeger, A. J., Shipley, T. F., & Reynolds, S. J. (2017). The roles of working memory and cognitive load in geoscience learning. Journal of Geoscience Education, 65(4), 506-518.
Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. American psychologist, 58(9), 697.
Kemps, E. (2001). Complexity effects in visuo-spatial working memory: Implications for the role of long-term memory. Memory, 9(1), 13-27.
Khodadadi, M., & Amani, H. (2018). Corsi spatial-visual working memory test [Software]. Institute for behavioral and cognitive sciences, Tehran, Iran. https://www.sinapsycho.com/Shop/Product/1319
Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. Educational psychology review, 21(1), 31-42.
Lai, E. R. (2011). Metacognition: A literature review. Always learning: Pearson research report, 24.
Larner, A. J. (2021). Functional Cognitive Disorders (FCD): How Is Metacognition Involved?. Brain Sciences, 11(8), 1082.
Manstead, A. S. R., Wagner, H. L., & MacDonald, C. J. (1983). A contrast effect in judgments of own emotional state. Motivation and Emotion, 7(3), 279-290.
Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). Metacognition: Knowing about knowing. MIT press.
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), 621.
Muncer, G., Higham, P., Gosling, C., Cortese, S., Wood-Downie, H., & Hadwin, J. A. (2021). A Meta-Analysis Investigating the Association Between Metacognition and Math Performance in Adolescence. Educational Psychology Review.
Nazari, M.A.& Nabizadeh Chianeh, Q. (2012). Checking the validity and reliability of Self-Assessment Manikin (SAM) Scale. Journal of Research in Psychological Health, 6(2).
Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 1–26). Cambridge, MA: MIT Press.
Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. Journal of educational psychology, 84(4), 429.
Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. Educational psychologist, 38(1), 1-4.

Pouget, A. JDrugowitsch, and A. Kepecs. 2016."Confidence and Certainty: Distinct Probabilistic Quantities for Different Goals.". Nature Neuroscience, 19(3), 366.
Rezapour, T., Ekhtiyari, H., & Soltaninejad, Z. (2013). procedure and scoring of Persian paper and pencil cognitive assessment package (PCAP). MEHRSA publication.
Ruiz, D., & Newell, A. (1989). Tower-noticing triggers strategy-change in the Tower of Hanoi: A Soar model. CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF PSYCHOLOGY.
Saed, O., Rushan, R., & Moradi, A. R. (2008). Investigating psychometric properties of Wechsler Memory Scale-for the students of Tehran Universities.
Sliwinski, M. J., Mogle, J. A., Hyun, J., Munoz, E., Smyth, J. M., & Lipton, R. B. (2018). Reliability and validity of ambulatory cognitive assessments. Assessment, 25(1), 14-30.
Sodian, B., Thoermer, C., Kristen, S., & Perst, H. (2012). Metacognition in infants and young children. Foundations of metacognition, 119-133.
Souza, A. S., & Oberauer, K. (2015). Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. Psychonomic Bulletin & Review, 22(1), 156-162.
Stanovich, K. E. (1999). Who is rational?: Studies of individual differences in reasoning. Psychology Press.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive science, 12(2), 257-285.
Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. Educational psychology review, 10(3), 251-296.
Sweller, J. (2011). Cognitive load theory. In Psychology of learning and motivation (Vol. 55, pp. 37-76). Academic Press
Turnbull, O. H., Denis, M., Mellet, E., Ghaëm, O., & Carey, D. P. (2012). The processing of visuo-spatial information: Neuropsychological and neuroimaging investigations. In Imagery, Language and Visuo-Spatial Thinking (pp. 97-124). Psychology Press.
Van Rooij, M. M., & Favela, L. H. (2016). A nonlinear dynamical systems theory perspective on dual-processing accounts of decision-making under uncertainty.
Wiley, J., & Jarosz, A. F. (2012). How working memory capacity affects problem solving. In Psychology of learning and motivation (Vol. 56, pp. 185-227). Academic Press.
Williams, M., & Woodman, G. F. (2012). Directed forgetting and directed remembering in visual working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(5), 1206.