The Effect of Mode of Presentation on Tower of Hanoi Problem Solving

Madison Fansher[1], Priti Shah[1], Sébastien Hélie[2]

[1]University of Michigan, [2]Purdue University

Corresponding Author

Madison Fansher,

530 Church St, Rm 1004

Ann Arbor, MI, 48109

Ph: 765-967-1225, E-mail: mfansher@umich.edu

Abstract

The Tower of Hanoi (TOH) is a classic problem that can be solved via multiple strategies. This study used TOH to examine how mode of presentation of a problem influences strategy use and transfer. Undergraduate students (Experiment 1) or Prolific workers (Experiment 2) completed two TOH problems of varying difficulty (4-disk/5-disk). They were randomly assigned to different conditions in which problems were either high in internal representation (mental) or high in external representation (computer). Participants were better able to complete problems successfully when external representations were available but completed problems in fewer moves when relying on internal representations. In addition, participants spent more time between moves when solving problems mentally, suggesting that external representations encourage speed while internal representations promote accuracy when solving recursion problems. Lastly, both experiments provide evidence that first solving a problem mentally encouraged participants to use strategies similar to goal recursion on a second problem.

Keywords: Problem Solving, Strategy Learning, Internal Representations, External Representations

The Effect of Mode of Presentation on Tower of Hanoi Problem Solving

To what extent do different types of representations impact one's ability to use and learn complex strategies when solving problems? One complex strategy often required for problem solving is goal recursion, a strategy where one must solve sub-goals or slightly simpler versions of the original problem to reach a solution (Simon, 1975). Components of goal recursion, including means-end analysis and establishing subgoals, are generalizable problem-solving strategies that are applicable to a wide variety of contexts. In educational contexts, goal recursion is most often required for mathematics and computer science problems (Pirolli & Anderson, 1985). It is reasonable to assume that facilitating the development of basic problem-solving strategies may support problem solving in a range of contexts both in and out of the classroom. One factor that may facilitate or impede the learning of such strategies are the internal and external representations available to the problem solver; external representations are physical depictions of a problem or constraints provided by the environment while internal representations include anything that is represented mentally such as ideas, strategies, and mental imagery (Zhang & Norman, 1994). The goal of this research was to assess whether individuals were more likely to use and learn complex strategies when Tower of Hanoi (TOH) problems were represented with an external visual representation (i.e., on a computer screen) or with an internal mental representation (i.e., mental imagery).

The Tower of Hanoi Problem

TOH is a classic, well-structured, and highly studied problem that has been used in cognitive science for over 100 years (e.g., Lucas & Claus, 1883). Cognitive scientists have long used TOH to understand human problem-solving processes (Klahr & Robinson, 1981; Kotovsky et al., 1985; Simon, 1975). The structure of TOH is what makes it useful. TOH is an example of a t*ransformation problem*, which are problems with an initial state, goal state, and rules associated with moving from one problem state to another (Greeno, 1978). In the standard 3-disk TOH puzzle, three disks of different sizes are stacked on a peg such that the largest disk is at the bottom and the smallest on top (See Figure 1). There are three total pegs, and the goal is to re-stack the disks on the third peg with the following constraints (1) only one disk can be moved at a time, (2) a larger disk cannot be placed on top of a smaller disk and (3) a disk cannot be moved if there is another disk on top of it. There are countless variations of the task (for example, with different start and end states, different numbers of disks, and isomorphic contexts).

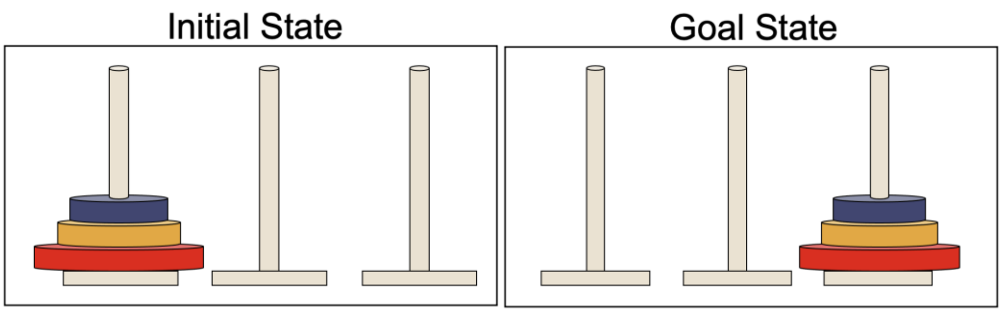

Figure 1. Shows the initial (left) and goal (right) states of the TOH problem.

Transformation problems are commonly solved using means-end analysis where the problem solver gradually reduces the difference between the initial and goal state (Anderson, 1993; Newell & Simon, 1972). TOH is used as a model task in the study of human problem solving because it is highly restricted and requires well-defined strategies for solving the problem such as goal recursion and remembering move patterns (Karat, 1982; Simon, 1975). For similar reasons, TOH has been used as a method for teaching the concept of recursion in computer science and mathematics (Ford, 1984) and has also been used as a measure of executive dysfunction in neuropsychology patients (Goel & Grafman, 1995).

Transfer of Strategy in Problem Solving

General heuristics and algorithms, such as goal recursion, setting sub-goals, and planning, can be applied to many problem-solving scenarios (Bassok & Novick, 2012; Gick, 1986). It is possible that once these general strategies are learned, that people will apply these strategies to a variety of problem-solving contexts (Adams, 1986). For example, learning goal recursion by solving TOH may help you solve a future problem requiring goal recursion, such as understanding factorials. Unfortunately, students often struggle to understand these general strategies and to apply them to novel problems. While there have been attempts to teach such general strategies to students, demonstrable transfer of these strategies to other tasks has been elusive (Barnett & Ceci, 2002; Catrambone & Holyoak, 1989; Klahr & Chen, 2011; Spencer & Weisberg, 1986)

Research finds that individuals can learn to solve problems that require more abstract strategies but are unable to solve isomorphic problems (i.e., structurally identical problems with a different surface structure) (Bassok & Holyoak, 1989; Hayes & Simon, 1977). For example, a classic TOH isomorph is the “Monsters and Globes” problem (Hayes & Simon, 1977), where three different sized monsters each hold a globe of a specific size, and the monsters must pass around the globes following a set of rules until each monster holds the globe corresponding to their size. One factor that may facilitate or impede developing an abstracted understanding of strategies is the extent to which they are learned via mental simulation or with the support of external visualizations. Solving visuospatial problems such as TOH without an external visualization requires strategic planning due to the large working memory demands of the problem (Handley et al., 2002; Zook et al., 2004). Internally using a trial-and-error strategy would be risky, as participants may lose track of the locations of the disks due to the many items that must be kept track of in working memory. As such, it is reasonable to predict that solving a problem using internal mental representations may facilitate learning and transfer of recursion, as strategic thought is required to successfully maintain a mental representation of the problem.

Distributed Representations of Problems

Problem solving is frequently characterized in terms of moving through a "problem space" that contains information about the problem, such as the current state of the problem (e.g., all the disks are on the leftmost peg), the goal, and possible actions one could perform. This information can be represented internally and/or externally. Internal representations include ideas, memories, and mental images, while external representations include physical images or objects, along with the constraints provided by the physical world. In a classic TOH problem in which one physically moves wooden disks from peg to peg, some information is externally represented (e.g., where the disks currently are; you cannot move a disk if there is another disk on top of it) whereas other information is internally represented (e.g., what the problem state will look like when the top disk is moved over one peg). Thus, the representation of most TOH problems is distributed, with some information being represented internally and other information represented externally (Zhang & Norman, 1994).

Internal and external representations guide, constrain, and even determine cognitive behavior (Zhang, 1997). The same problem can be designed to provide more or less external support and/or to alter the internal cognitive processing demands. One method for manipulating external and internal demands in the TOH is varying mode of presentation of the problem. For example, individuals could solve the problem entirely mentally or with an external representation that is always available.

Internal and external representations have different problem-solving affordances. In general, external representations offload cognitive demand, turning a difficult cognitive process into a simpler perceptual inference (Kotovsky et al., 1985; Larkin & Simon, 1987; Wu & Shah, 2004). When mentally simulating TOH, working memory resources are taxed by the need to keep track of the current problem state; when the problem is presented externally the current state is always visible. According to Cognitive Load Theory, the requirement to hold the current state in working memory may be construed as extraneous and leave the solver with fewer precious cognitive resources needed for forming appropriate problem representations and planning (Sweller, 1988).

Consistent with Cognitive Load Theory, some researchers have found that using external memory aids, such as pencil and paper, increases problem solving performance by reducing cognitive load (Cary & Carlson, 2001). Likewise, Barrett, Stull, Hsu, and Hegarty (2015) found that in the domain of organic chemistry, people were able to offload cognitive demands onto a visual display of a chemistry problem. In addition to the benefits from offloading working memory, Bocanegra, Poletiek, Ftitache, and Clark (2019) argue that being able to see problems visually allows the solver to reconceptualize the problem and interact with problems in a more meaningful way. Others suggest that technological aids may increase strategic thinking as extraneous details are offloaded onto the aid, leaving more processing resources for complex thought (for a review, see Reiser, 2004).

Alternatively, it is reasonable to question whether solving TOH problems with an available external representation would promote passive problem solving, especially when individuals are not highly motivated to think strategically (Schoenfeld, 1991). The external representation acts as an extension of working memory, and thus, people may successfully complete problems through trial-and-error because there is little cognitive resource cost associated with making a move (Hélie & Pizlo, in press). Thus, developing and using an efficient strategy is not necessary. In contrast, there is a greater cost to updating and keeping track of the current problem state when a TOH problem is solved mentally, and a trial-and-error strategy is likely to overwhelm working memory resources. Therefore, individuals who are required to solve TOH problems mentally may be more frugal with the number of moves they use.

Noyes and Garland (2003) provide an initial test of this hypothesis. Across three within-subject design experiments, participants solved TOH problems mentally, with a physical TOH, and with a computer. The authors found modest evidence that participants solved problems with fewer moves in the mental condition. However, this effect was not consistent. They consistently found that participants spent more time between moves when solving problems mentally which led the authors to suggest that participants engaged in more planning and implemented a more complex strategy than participants with an external representation. They also found that participants were consistently more likely to finish TOH problems in the computer condition in comparison to the mental condition.

The Current Study

The current study expands on Noyes and Garland's work by thoroughly outlining a protocol for studying internal representation-based problem solving and addressing how problem difficulty (e.g., number of disks) and representation interact to influence strategy selection. It also explores whether internal representations promote the development of generalizable strategies that transfer to more difficult problems. In each experiment, participants solved TOH problems with high external (on a computer) or internal (mentally) representation. An initial pilot study was conducted to gauge the impact of number of disks on problem solving performance, allowing for the definition of easy and difficult problems that could be solved within a 15 min time frame by most participants (see Supplementary Materials). In the main text, we report the results of two experiments examining the impact of problem representation on strategy use and transfer, where Experiment 2 serves as a pre-registered replication of Experiment 1's main findings in a different sample. In both experiments, participants were randomly assigned to one of four conditions: MC (Mental/Computer), CM (Computer/Mental), CC (Computer/Computer), or MM (Mental/Mental), in which the first letter indicates the mode of presentation of the first problem and the second letter indicates the mode of presentation of the second problem.

Experiments 1 and 2 examined how initial and subsequent problem representation influenced performance on an easy TOH problem (4-disk), followed by a more difficult problem (5-disk). Performance measures were completion rate, progression toward the goal state, number of moves, and time between moves. We predicted that participants would be more likely to successfully finish problems on the computer, but participants solving problems mentally would complete them using fewer moves. Additionally, we predicted that individuals who solved the initial problem mentally would solve a second problem of greater difficulty using fewer moves than those who solved the first problem on the computer. To anticipate, we found that participants were more likely to finish difficult problems when given an external representation but were more likely to use an efficient strategy when solving difficult problems mentally, as assessed by performance and self-reported strategy use. We also found modest evidence that solving a first problem mentally facilitated performance on a more difficult problem. All data and analysis scripts are available at https://osf.io/puqs2/?view_only=81b4b4a7c59040c0a35d808ae299a4cc,

Experiment 1

The overall goal of this study was to examine how internal and external representations affect problem solving. When people learn to solve problems, some may spontaneously learn generalizable strategies that allow for completing isomorphic problems or problems of greater or lesser difficulty. Experiment 1 examined how internal and external representation would influence complex strategy use and transfer to a subsequent problem of greater difficulty. Participants solved an easy (4-disk) followed by a difficult (5-disk) TOH problem, with mode of presentation of the problems varying by condition (MC, CM, MM, CC; M = Mental, C = Computer). It was deemed appropriate to use a 4-disk problem as an “easy” problem and a 5-disk problem as a “difficult” problem based on pilot data (see Supplementary Materials). Performance for each problem was assessed based on completion of the TOH, progress toward the goal state, the number of moves taken to complete the problem, and the average time between moves. Based on findings from Noyes and Garland (2003), it was hypothesized that solving problems with a computer would facilitate completion but not the use of complex strategies, where strategy use is inferred when participants complete the problem in fewer moves and with more time spent between moves.

Strategy transfer between modes of presentation was examined with linear regression models including the mode of presentation of the first and second problems viewed by participants as predictors. This method allowed for examining how viewing the initial problem in one mode of presentation would influence performance on a second problem. Experiment 1 consisted of three components: (1) solving a 4-disk TOH, (2) solving a 5-disk TOH, and (3) completing a post-test questionnaire. The results suggest that participants are more likely to *solve* the TOH problem when they rely on external representations (computer) but that they solve the problem more *strategically* when they rely on internal representations (mental).

Methods

Participants

One hundred thirty-eight students were recruited from the undergraduate population of Purdue University and received partial fulfillment of a course requirement for their participation. A sample size of at least 30 participants per condition was chosen out of convenience. Participants were ineligible to participate if they were familiar with TOH. All participants were randomly assigned to one of four conditions (MC = 33, CM = 33, MM = 35, CC = 37). All procedures were approved by the Purdue University Institutional Review Board.

Materials

Computer Conditions. In conditions where participants solved problems on the computer (CC, CM, and MC), participants interacted with a virtual TOH programmed with Python 3. The problem display contained disks labelled 1 to *n*, with *n* representing the total number of disks, and three pegs labelled “A”, “B”, and “C” (see Figure 2).

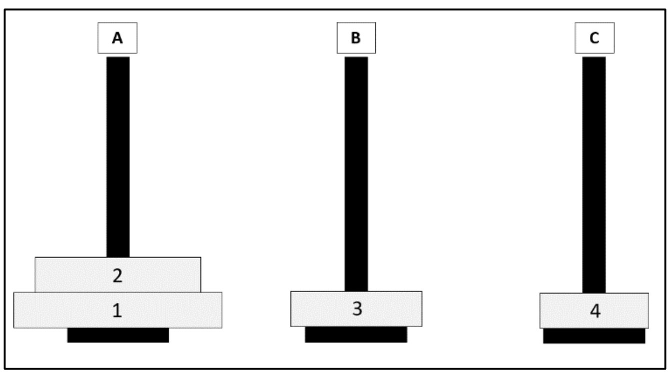

Figure 2. Example virtual TOH display with four disks

All responses were recorded with key presses on the number pad of a standard QWERTY keyboard with an “A” label on the “7” key, a “B” label on the “8” key, and a “C” label on the “9” key. To move a disk, participants pressed the key on the number pad corresponding to the disk, followed by the key with the label of the peg they wanted to move the disk to. For example, if one wanted to move disk 4 to peg B, they would press button 4 followed by button 8 (which has a “B” sticker label on it).

Participants were shown an error message if they violated one of the problem rules or pressed an incorrect key. The error messages included “You cannot move a disk with another disk on top of it”, “You cannot put a larger disk on top of a smaller disk”, and “The disk is already on the pole you selected". For each problem, we recorded all moves and the time between each move, total time to completion, total number of moves, the number of times a rule was broken (e.g., attempted to move a larger disk on top of a smaller disk), and an image of what the problem space looked like when the participant either ran out of time or completed the problem.

Mental Conditions. Participants solving problems mentally (MM, CM, MC) were given a static image of the problem space on a piece of paper to use as a visual aid. Participants in these conditions mentally kept track of the disks and verbally reported their moves to the experimenter (e.g., “move disk 1 to peg C”). The experimenter recorded the participant’s progress by completing the computer program described above with the participant’s instructions. Participants were unable to view the experimenter’s screen while solving the problem, unless they reported to the experimenter that they could no longer remember where the disks were located (see Procedure). Excluding the verbal protocol required in the mental condition (see Discussion), this method was identical to the procedure used by participants solving problems on the computer. Both conditions required that the participant select a disk followed by a peg, whether that be with two keyboard presses or by verbally reporting to the experimenter. After completing the TOH problems, participants completed an online questionnaire regarding their perceived difficulty of the problems and the strategies that they used while solving the problems (see the Appendix for specific questions).

Design

The experiment implemented a 4 × 2 mixed design with condition as a between-subjects factor with four levels (MC, CM, CC, and MM) and problem difficulty as a within-subject factor (4-disk, 5-disk). Performance was assessed with four dependent variables: problem completion, progress toward the goal state, number of moves, and average time between moves.

Procedure

Instructions. Participants were randomly assigned to MC, CM, MM and CC conditions with a random number generator. After being screened for prior experience with the problem and providing informed consent, participants were told that they would solve a series of problems and that they would have 15 min to complete each problem. The experimenter showed all participants an example 3-disk TOH and the problem rules were explained. Participants were given instructions on how to solve the problem using the computer program and with the mental protocol across all conditions. They watched the experimenter make one move and one mistake with the computer program to better understand the procedure. They were also shown the visual aid that would be provided if they were to solve a problem mentally. Participants were informed that they would be allowed to briefly look at the experimenter’s screen if they were solving a problem mentally and could not remember where the disks were located. However, they were also told they should do this as little as possible. Participants had the opportunity to ask any questions and were then given a 4-disk problem which was determined to be of low difficulty based on findings from a pilot study. After 15 min (or problem completion, whichever came first), participants were told the mode of presentation for the subsequent 5-disk problem. After 15 min, or problem completion, participants were given a posttest questionnaire.

Computer Procedure. Participants who solved problems with a computer (MC, CM, CC) interacted with the TOH program described in the Materials.

Mental Procedure. Participants who solved problems mentally (CM, MC, MM) were told to imagine moving the disks from peg to peg and that they could use the visual aid as needed; they were not allowed to write on the visual aid. They verbally reported each move to the experimenter, who solved the problem with the computer program, following the participant's instructions. The experimenter and participant sat side-by-side and were separated by a cubicle divider blocking the participant's view of the experimenter's screen.

During the problem, the experimenter would inform the participant if they broke a problem rule and the participant informed the experimenter whenever they needed to be reminded of the location of the disks. The experimenter kept track of how many times participants asked to look at the screen for each problem. It was stressed to participants that the task was a mental task and that they should try to look at the screen as little as possible. This mode of presentation relied heavily on internal representation but allowed participants to be periodically reminded of the state of the problem space.

Results

Exclusion Criteria

Data were excluded from 3 participants for not understanding the TOH rules (n = 135). In addition, a computer malfunction led us to completely remove 11 additional participants from the data analysis (*final* n = 129).

Performance on the first problem (4-disk)

Performance on the 4-disk problem was assessed with a series of Welch two-sample t-tests comparing performance in computer conditions (CC and CM) to performance in mental conditions (MM and MC), unless otherwise noted.

Completion rate. A Chi-Square test compared the proportion of participants who successfully solved the problem in the 15-min time period between conditions. Participants were more likely to complete the problem with the computer (100%) than the mental (91%) mode of presentation ($\chi^2(1) = 4.45$, $p = .03$).

Progress to the goal state. In addition to assessing problem completion as a binary outcome, a progress score was calculated to compare performance on a continuous scale. Progress was calculated by adding the number of disks in the correct order on the last peg (considered moving toward the goal state) and subtracting the number of disks in the incorrect order on the last peg (considered moving away from the goal state). For example, someone with all four disks on the last peg receives a perfect score of 4, while someone with the largest and smallest of the four disks receives a score of 0 (+1 for having the largest disk on the last peg on the bottom, -1 for having the smallest disk on the last peg in the incorrect order).

As such, the computer condition made more progress ($M = 4.00$) than the mental condition ($M = 3.61$), as expected considering all participants in the computer condition completed the problem ($t(35) = 2.44$, $CI_{95\%} = [.07, .71]$, $p = .02$).

Number of moves. One way to assess strategy use is through the number of moves it takes participants to solve the problem. The closer their score is to the optimal score (in this case, 15 moves), the more likely it is that participants are using a strategy that is more complex than trial-and-error. This analysis only includes participants who completed the problem within the 15-minute time period (Mental ($n = 58$), Computer ($n = 65$)) (Fig 3A illustrates number of moves for all participants). Normalized number of move scores were computed for each problem with the following metric: $normalized\ number\ of\ moves =$ $\frac{number\ of\ moves - optimal\ number\ of\ moves}{optimal\ number\ of\ moves}$ where the number of optimal moves for 4 and 5-disk problems are 15 and 31 moves, respectively. After normalizing the number of moves, a score of 0 indicates optimal performance, while the larger the score the less optimal. There was no statistically significant difference in number of moves between conditions for the 4-disk problem ($t(116.39) = 1.14$, $CI_{95\%} = [-.10, .37]$, $p = .26$).

Average time between moves. While the computer and mental conditions are virtually parallel in respect to how participants make their responses (i.e., "4 to C"), they differ on one critical element: the verbalization protocol. If participants are slower in the mental condition, this isn't informative given that they must articulate their moves to the experimenter, while the computer condition moves the disks by key press. As an effort to correct this confound, for the mental condition, the estimated time it would take to articulate a move was subtracted from the average time between moves.

Jacewicz et al. (2009) found that the articulation rate for young adults when engaging in spontaneous speech (i.e., not reading) is approximately 5.18 syllables per second. In the current study, when reporting moves to the experimenter, participants would engage in 3-4 syllable utterances (i.e., "move 4 to C", or "4 to C"). According to Jacewicz et al. (2009) it would take participants ~.77 sec to articulate a four-syllable utterance. Thus, to calculate the time between moves for each participant we divided the total amount of time to spent on the problem by the total number of moves, and for those in the mental conditions we subtracted .77 sec from this value. Even after this correction, participants in the mental condition ($M = 14.04$ sec) spent significantly more time between moves than those in the computer condition ($M = 7.58$ sec) ($t(75.99) = -5.62$, $CI_{95\%} = [-8.76, -4.20]$, $p < .001$), with this difference being much larger than the correction itself.

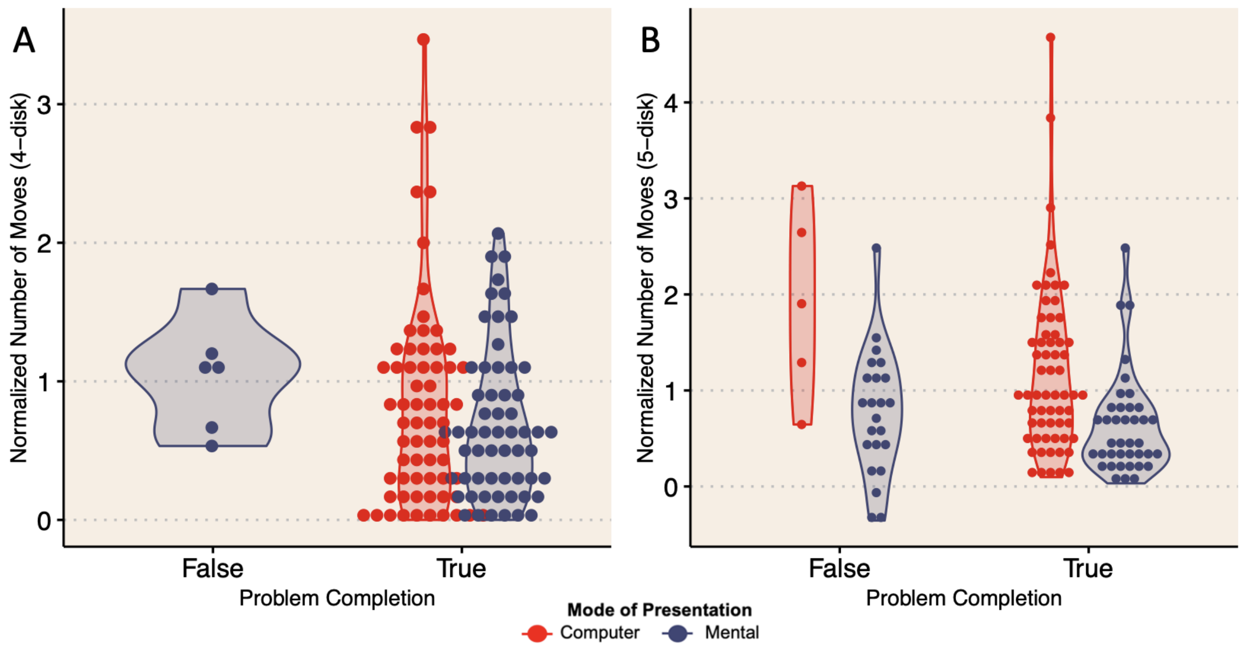

Figure 3. *Number of moves for those who did (right) and did not (left) complete the 4-disk (A) and 5-disk (B) TOH problems in the allotted 15 min*. These data illustrate that the effect of condition on number of moves is not driven by a selection effect in which only the best participants in the mental condition were retained for the analysis.

Performance on the second problem (5-disk)

Performance on the second, 5-disk problem, was assessed with a series of linear regression models with either binary (i.e., completion rate), or continuous outcome variables (i.e., progress to goal state, number of moves, and time between moves), unless otherwise noted. Each model includes mode of presentation on the first, 4-disk problem, mode of presentation of the second, 5-disk problem, and their interaction as predictors.

Number of reminders. As mentioned in the Methods section, participants who solved problems mentally were allowed to ask the experimenter to see the state of the problem if they could no longer remember where the disks were. The experimenter recorded how many times each participant asked to look at the screen. Before further analysis, the number of reminders were compared between the MM and CM groups to ensure there were no confounds introduced to the analyses since the groups would be compared to one another on other outcome variables. Number of reminders was not compared between MM and MC participants in the 4-disk analysis because MC and MM participants are never directly compared to one another (i.e., we collapse across mental and computer conditions when examining performance on the 4-disk problem). A Welch two-sample t-test indicated no difference between the number of times MM (M = 2.00) and CM (M = 2.13) participants looked at the screen (t(53.95) = .28, $CI_{95\%}$ = [-.82, 1.08], p = .78).

Completion rate. Mode of presentation of the second problem significantly predicted whether participants completed the 5-disk problem in the 15-minute time period (B = -2.67, $SE(B)$ = .82, p = .001), as more participants completed the problem in the computer (93%) than mental (63%) conditions. There was no main effect of the mode of presentation of the first problem (B = -0.53, $SE(B)$ = .95, p = .57) nor a significant interaction between mode of presentation for the first and second problems (B = 1.34, $SE(B)$ = 1.09, p = .22) (see Fig 4A).

Progress to the goal state. Mode of presentation on the second problem significantly predicted progress toward the goal state, as participants in the computer condition (M = 4.58) were overall closer to the goal state than those in the mental condition (M = 2.90) (B = -2.55, $SE(B)$ = .56, t(125) = -4.58, p < .001). There was no main effect of mode of presentation on the first problem (B = -.22, $SE(B)$ = .55, t(125) = -.40, p = .69). However, there was a significant interaction between mode of presentation for the first and second problems (B = 1.71, $SE(B)$ = .79, t(125) = 2.16, p = .03). Participants who solved the second problem with the mental mode of presentation made significantly more progress if they had first solved a problem mentally, as opposed to on the computer.

Number of moves. This analysis only includes participants who successfully completed the 5-disk problem in the 15-minute time period (Mental (n = 39), Computer (n = 62)); for a distribution of all the data, see Figure 3B. Mode of presentation on the second problem significantly predicted normalized score on the 5-disk problem (B = -.47, $SE(B)$ = .23, t(97) = -2.04, p = .04), as participants in the mental condition solved the problem in fewer moves when compared to the computer condition (see Figure 4B). There was no effect of mode of presentation of the first problem (B = .16, $SE(B)$ = .19, t(97) = .81, p = .42), nor a significant interaction between the first and second problem mode of presentation (B = -.17, $SE(B)$ = .31, t(97) = -.53, p = .59).

Average time between moves. Average time between moves was again corrected for articulation in the mental condition and there was a significant effect of mode of presentation on the second problem, such that participants in the mental condition spent more time between moves than those in the computer condition (B = 8.91, $SE(B)$ = 1.37, t(125) = 6.52, p < .001). There was no main effect of mode of presentation of the first problem (B = 1.11, $SE(B)$ = 1.34, t(125) = .82, p = .41) but critically, there was a significant interaction between the first and second problem mode of presentation (B = -4.57, $SE(B)$ = 1.93, t(125) = -2.36, p = .02). Participants who solved the second problem on the computer spent *more* time between moves if they had first solved a problem mentally, while participants who solved the second problem mentally spent *less* time between moves on the second problem if they had first solved a problem mentally (see Figure 4C).

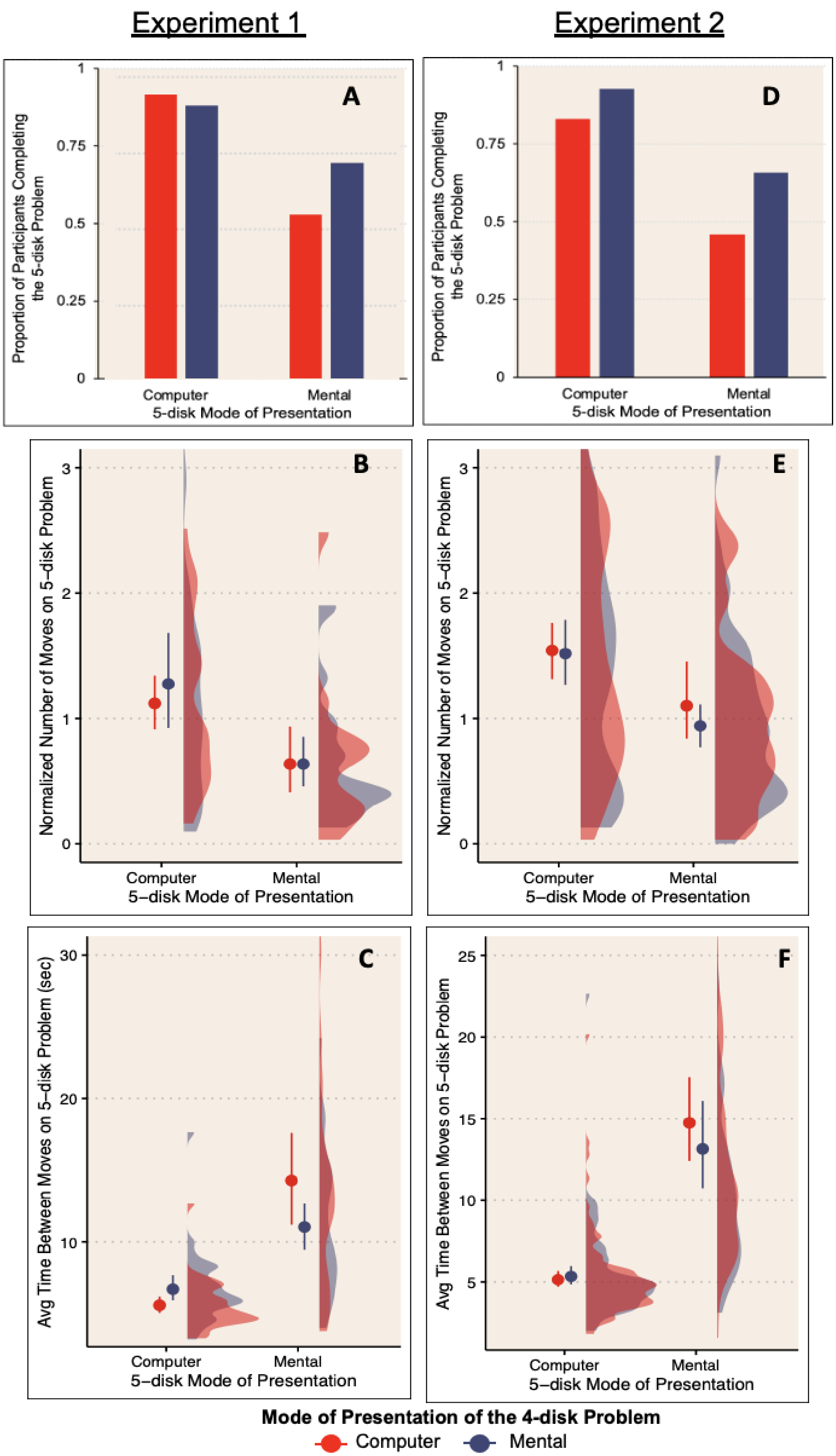
Experiment 1
Experiment 2
A
D
Proportion of Participants Completing the 5-disk Problem
0
0.25
0.5
0.75
1
Computer
Mental
5-disk Mode of Presentation
B
E
Normalized Number of Moves on 5–disk Problem
0
1
2
3
Computer
Mental
5–disk Mode of Presentation
C
Avg Time Between Moves on 5–disk Problem (sec)
10
20
30
F
Avg Time Between Moves on 5–disk Problem
5
10
15
20
25
Mode of Presentation of the 4-disk Problem
Computer
Mental

Figure 4. A – C illustrate the proportion of participants completing the 5-disk problem in each group, normalized number of moves on the 5-disk problem by group, and average time between moves on the 5-disk problem by group. D-E illustrate the same data but for Experiment 2. In each figure, mode of presentation of the 5-disk problem is presented on the x-axis and mode of presentation of the first 4-disk problem is indicated by color, with C shown as red and M shown as dark blue. Error bars illustrate 95% confidence intervals.

Self-Reported Strategy Use

Explicit strategy use was examined with an exploratory analysis of the strategies self-reported by participants (see Appendix). For Experiment 1, six additional participants failed to correctly submit the questionnaire leaving 123 observations (CC = 35, CM = 28, MM = 30, MC = 30).

In the questionnaire, participants were asked to indicate the percentage of the time that they used the following strategies: *remembering move sequences*, *using sub-goals*, *developing step-by-step instructions to solve the problem*, *trial-and-error*, and *no strategy*. The strategies were selected based on the various strategies outlined by Simon (1975) and the factor analysis completed by Noyes and Garland (2003) that extracted similar components from the self-reported strategies of participants solving TOH.

Participants were asked to report the percent of time that they used each strategy on a scale of 0-100% using a continuous slider bar interface. Participants in conditions where they viewed both problems in the same mode of presentation completed this task once (CC & MM), participants who viewed the problems in different modes of presentation did this twice, reporting the strategies they used with the mental mode of presentation and with the computer mode of presentation separately (CM & MC).

Since participants reported each strategy on a scale of 0-100%, the total percentage of strategies used often sums to be over 100. Percentage of the time participants used each strategy overall was calculated by dividing their slider-scale response for each strategy by the sum of all strategy percentages reported. Significance tests that are not reported in the manuscript, as well as additional analyses on the self-report data, are available in the Supplementary Materials.

Same mode of presentation groups. A factorial ANOVA examined strategy use between participants who only solved problems on the computer (CC) and participants who only solved problems mentally (MM). In this analysis, mode of presentation and the type of strategy were included as predictors with percentage of strategy use as the dependent variable. There was a significant main effect of strategy ($F(4, 315) = 9.54$, $p < .001$). Post-hoc tests show that, overall, participants reported using *no strategy* less than all of the other strategies, and also used *remembering move sequences* more than *trial and error*. There was also a significant interaction between mode of presentation and strategy ($F(4, 315) = 2.93$, $p = .02$) and this interaction was driven by MM participants self-reporting *remembering move sequences* more ($M = .34$) than CC participants ($M = .22$) (see Figure 5). There was no main effect of mode of presentation ($F(1,315) = 0$, $p = 1$.)

Different mode of presentation groups. For participants who solved problems both on the computer and mentally, strategy use for the two modes of presentation were compared within-subject. A series of factorial ANOVAs are reported with mode of presentation and condition (CM or MC) as the predictors and percentage of strategy use as the dependent variable.

Participants reported using *remembering move sequences* ($F(1,112) = 33.86$, $p < .001$) and using sub-goals ($F(1,112) = 4.46$, $p = .04$) more when they solved the problem mentally compared to on the computer. Participants also reported using *trial and error* more when solving problems on the computer than when solving problems mentally ($F(1,112) = 5.51$, $p = .02$). No other differences in strategy use within subject were significant (see Figure 5).

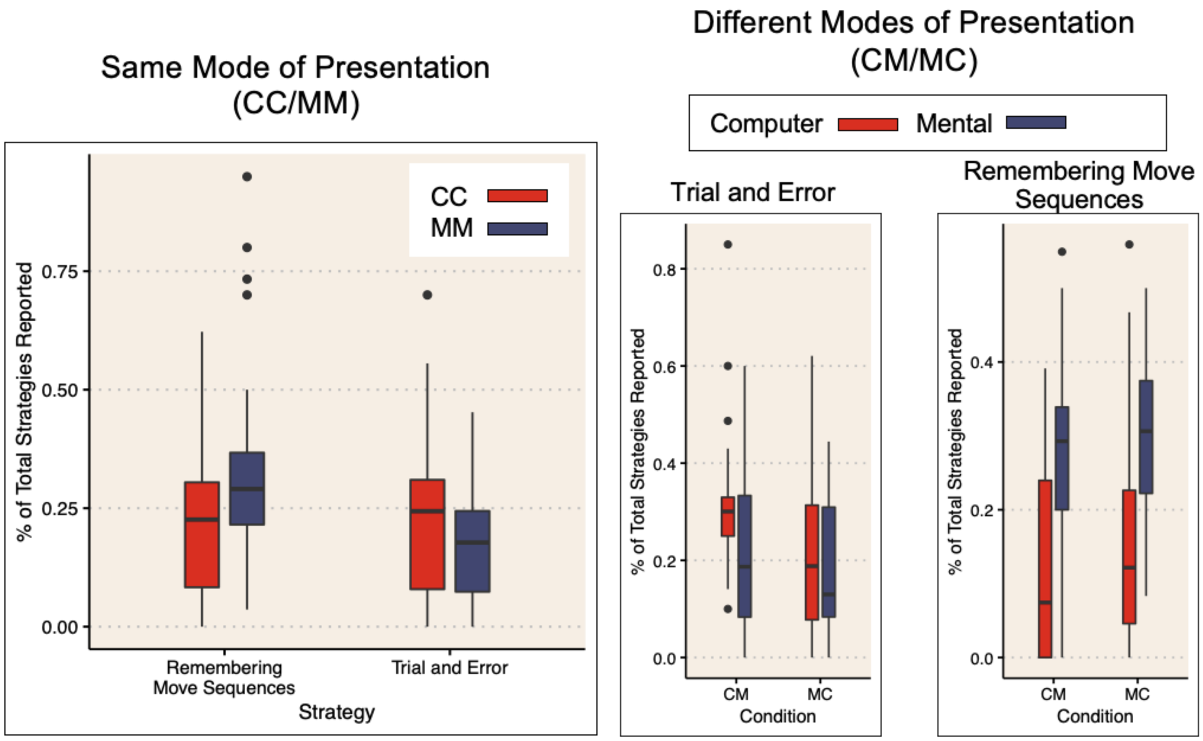

Figure 5. The left panel illustrates strategy use as self-reported by participants who viewed both problems with the same mode of presentation. The right panel compares the use of different strategies within-subject for participants who solved problems with multiple modes of presentation.

Discussion

Overall, the results in Experiment 1 suggest that having an external representation allowed participants to complete the TOH problems within the allotted time. Participants in the computer condition were more likely to complete the problems and made more progress toward the goal state regardless of difficulty. There was also evidence suggesting that participants in the mental conditions engaged in more complex strategies, as participants spent more time between moves regardless of difficulty and solved the difficult (5-disk) problem in fewer moves than those in the computer condition. One might assume that asking the experimenter for reminders of the problem state led to the difference in time observed between moves computer and mental conditions. However, the number of times participants asked for reminders in mental conditions was relatively low across experiments. The median number of reminders was 1 for both the 4- and 5-disk problems. The strategy self-report findings corroborate these assumptions, as both within- and between-subject analyses suggested that participants used *remembering move sequences* (akin to goal recursion) and *sub-goals* more when solving problems mentally, and *trial and error* more when solving problems on the computer.

In addition, there was modest evidence supporting the notion that participants who first solved a problem mentally used a different strategy when solving a subsequent problem than those who first solved a problem on the computer. This was illustrated by the interaction between the first and second problem mode of presentation on progress toward the goal state and average time between moves on the 5-disk problem. Participants who solved the 5-disk problem mentally made more progress toward the goal state if they had first solved the 4-disk problem mentally in comparison to on the computer. In addition, participants who solved the 5-disk problem mentally spent less time between moves if they had first solved the 4-disk problem mentally in comparison to on the computer. The opposite pattern held true for those who solved the 5-disk problem on the computer, that is, first solving the 4-disk problem mentally led to longer time between moves when compared to participants who first solved the problem on the computer. It is possible that these interactions result from participants developing better strategies when solving the 4-disk problem mentally, or these interactions could simply be a result of transfer appropriate processing, where participants perform better on a second problem that is in the same mode of presentation as the first problem they solved (see General Discussion for more detail).

The purpose of Experiment 2 was to replicate this finding in a larger, more diverse sample, given that our small sample sizes in Experiment 1 may have prevented additional transfer effects from emerging. In addition, the design of the mental condition was updated so that participants in both computer and mental conditions recorded their moves with keypresses, thus eliminating the verbalization component of the mental procedure.

Experiment 2

Experiment 1 examined how mode of presentation affected performance and strategy transfer from an easy to a difficult TOH problem. Experiment 2 sought to replicate the main findings with a larger and more diverse sample. In addition, the mental condition was redesigned so that participants did not have to orally communicate their moves to the experimenter. Rather, they entered their moves using the keyboard identically to participants in the external condition. This change removes the possible confound in Experiment 1 in which response modality differed for the internal (oral) and external (typing) conditions. Thus, the longer time per move for the mental condition could have been due to articulation time rather than planning. It also removes the possibility to that any strategy learning exhibited by the mental condition was due to verbalization.

Based on findings from Experiment 1, we hypothesized that participants would be more likely to complete the 4- and 5-disk problems and would make more progress toward the goal state when solving problems on computer. We also hypothesized that participants would spend more time between moves when solving problems mentally as compared to on the computer.

Lastly, we hypothesized that participants would solve the 5-disk problem in fewer moves when the problem was presented mentally[1]. Experiment 2 was pre-registered at osf.io/cjf6y

Methods

Participants

Five hundred seventy-two young adults were recruited from Prolific, an online platform for recruiting participants (Age *M(SD)* = 22.35(3.24); Sex = 65% F, 33.2% M, 1.7% Other). Young adults were recruited specifically as this was the demographic recruited for Experiment 1. In addition, the mental problems may have been too difficult for older adults to complete in the allotted time given that it is well-documented that visuospatial working memory declines with age (Reuter-Lorenz et al., 2000; Rowe et al., 2008) and that there are age-related differences in TOH performance (Rönnlund et al., 2010). Participants were paid $5.60 for their time based on Prolific's recommended rate for a ~30 min study. After data collection it is estimated that participants were paid, on average, $11.20/hr. All procedures were determined to be exempt by the University of Michigan IRB.

Note that the number of participants recruited deviates from our pre-registered sample size. Initially sample size was determined with an 80% power analysis using the size of the effect of mode of presentation on the number of moves to complete the 5-disk problem from Experiment 1 (f^2 = .05). This analysis suggested that data should be collected from 218 participants. To ensure that there would be sufficient power for the number of move analysis, where participants who do not complete the problems are excluded, the possibility that 40% of participants in the mental groups may not finish the 5-disk problem (~63% completion rate in Experiment 1) was accounted for. Thus, 304 participants total, or ~76 per group were recruited.

[1] It was also hypothesized that we would replicate the finding that participants self-report using more generalizable strategies in the mental than computer conditions but we do not report these data (see Methods).

However, it was not anticipated that technical difficulties would result in incomplete data for many participants, that a high percentage of participants would report prior experience with TOH problems, and that many people would report putting little effort into the study. After collecting data from 304 participants and applying these exclusion criteria ~30-40 participants remained per condition. Thus, data were collected from an additional 268 participants in hopes that at least 76 participants would remain per group after applying these exclusion criteria to the new data (see Results section).

Materials

Computer Conditions. In conditions where participants solved problems on the computer (CC, CM, and MC), participants interacted with a virtual TOH programmed with PsychoPy and integrated with Pavlovia, an online platform where researchers can host behavioral experiments. The problem display again contained disks labelled 1 to *n*, with *n* representing the total number of disks, however, the three pegs were now labelled “J”, “K”, and “L”. In Experiment 1 participants pressed keys on the number pad of a keyboard. In Experiment 2, to ensure that participants could respond on computer devices without a number pad, the labels of the pegs were changed to be “J”, “K”, and “L”. These keys were selected because they are next to each other and are farthest away from the row of numbers on a QWERTY keyboard. To move a disk, participants pressed the key in the row of numbers corresponding to the disk they wished to move, followed by the key with the label of the peg they wanted to move the disk to.

Participants were shown an error message if they violated one of the problem rules or pressed an incorrect key. The error messages included “You cannot move a disk with another disk on top of it”, “You cannot put a larger disk on top of a smaller disk”, and “The disk is already on the pole you selected". For each problem, all keypresses and time between keypresses, as well as the mode of presentation for each problem presented was recorded.

Mental Conditions. Participants solving problems mentally (MM, CM, MC) were shown a labelled static image of the initial and goal states of the problem space (akin to Fig. 1) on the computer screen to use as a visual aid. Participants in these conditions mentally kept track of the disks and recorded their moves with the same keypresses as participants in the computer condition. For example, someone in the mental condition would move disk “4” to peg “L” with two keypresses, however, the image on the screen would not update to reflect this change (it would still be a static image of the initial and goal problem states). If a participant made an error, they were shown the same error message as participants were shown in the computer condition. If participants completely forgot the state of the problem, they were instructed to press the “z” key on the keyboard. Doing so allowed participants to view the current state of the problem for 10 sec; participants could not move any of the disks during this time. They were encouraged to press the “z” key only when necessary.

Strategy Survey. Participants reported the amount of time they spent using the strategies different described in Experiment 1 when they were solving problems on the computer or mentally, depending on condition. Unfortunately, many of the participants were not redirected to the self-report survey at the end of Experiment 2 or failed to correctly indicate the mode of presentation in which they solved the problems. Failing to correctly self-report condition led participants to answer questions that were not relevant to their actual experience. We believe participants failed to correctly self-report their mode of presentation because they confused the practice and experimental problems. Thus, we do not analyze, or report results from the self-report data.

Design

The experiment implemented a 4 × 2 mixed design with condition as a between-subjects factor (MC, CM, CC, and MM) and problem difficulty as a within-subject factor with two levels. The same outcome variables were analyzed as in Experiment 1.

Procedure

Prolific workers were invited to participate in an experiment on complex problem solving. Those who were eligible to participate in the experiment clicked on a link which redirected them from Prolific to a consent form displayed with Qualtrics survey software. Those who consented were redirected to Pavlovia, which hosted the TOH program programmed with PsychoPy. Since the experiment was run online and an experimenter would not be present, participants were given extensive instructions on how to work the program, what the TOH rules were, and how to solve the problems visually or mentally (see OSF for instruction stimuli). Participants were told that they would solve two practice problems. They first solved a 3-disk problem visually so that they could practice moving disks around with their keyboard. After problem completion or 15-min had elapsed, they were instructed to solve another 3-disk problem mentally. After participants had finished the problem or 15-min had elapsed, they were told that they would start the actual experiment and that they would be told the mode of presentation for subsequent problems at the start of each problem.

Participants were evenly randomly assigned to CC, MM, CM, or MC conditions and viewed the 4-disk problem followed by the 5-disk problem in the modes of presentation associated with their condition. After 15 min had elapsed or participants had completed the problem, they would move onto the next problem. At the end of the 5-disk problem they reported prior experience with TOH, and indicated the amount of effort they put forth in the experiment.

At the end of the survey, they were given a completion code to enter into Pavlovia to confirm their participation for compensation.

Results

The same analysis methods were used to assess performance on the first and second problems as outlined in Experiment 1, in which performance on the first 4-disk problem was compared between modes of presentation (C or M) and was assessed with Welch two-sample t-tests (i.e., progress to the goal state, avg time between moves, normalized number of moves) or Chi-square tests (i.e., completion rate). Performance on the second, 5-disk problem was assessed with a series of regression models with mode of presentation of the 4-disk problem, mode of presentation of the 5-disk problem, and their interaction as predictors with either continuous (i.e., progress to the goal state, avg time between moves, normalized number of moves) or binary (i.e., completion rate) outcomes.

Exclusion Criteria

Data were excluded from 34 participants who had incomplete data due to user error. Data from 153 of the remaining participants were excluded as they had self-reported prior experience with TOH problems. From the remaining participants, data were excluded from an additional 14 participants who reported putting little effort in the study. After applying these pre-registered exclusion criteria, 371 participants remained with 106 in the CC group, 96 in the CM group, 84 in the MC group, and 85 in the MM group.

Performance on the first problem (4-disk)

Completion rate. Participants were more likely to finish the 4-disk problem in the computer group (95.54%) than the mental group (87.57%) ($\chi 2(1) = 6.83$, $p = .009$).

Progress to the goal state. There was no statistical difference in progress to the goal state when comparing those who solved the 4-disk problem on the computer (*M* = 3.73) to those who solved the problem mentally (*M* = 3.60) ($t(240.83) = .61$, $CI_{95\%} = [-.29, .54]$, $p = .54$).

Number of moves. This analysis only included participants who completed the 4-disk problem. For a distribution of all the data see Figure 6A. Normalized number of moves did not differ between those in mental (*M* = 1.04) and computer (*M* = 1.11) conditions ($t(329.69) = .62$, $CI_{95\%} = [-.15, .29]$, $p = .54$).

Average time between moves. In contrast to Experiment 1, correcting for verbalization in the mental condition was not necessary. As hypothesized, participants took significantly more time between moves on average when solving the 4-disk problem mentally (*M* = 12.69 sec sec) compared to on the computer (*M* = 5.51 sec) ($t(196.43) = -9.88$, $CI_{95\%} = [-8.61, -5.74]$, $p < .001$).

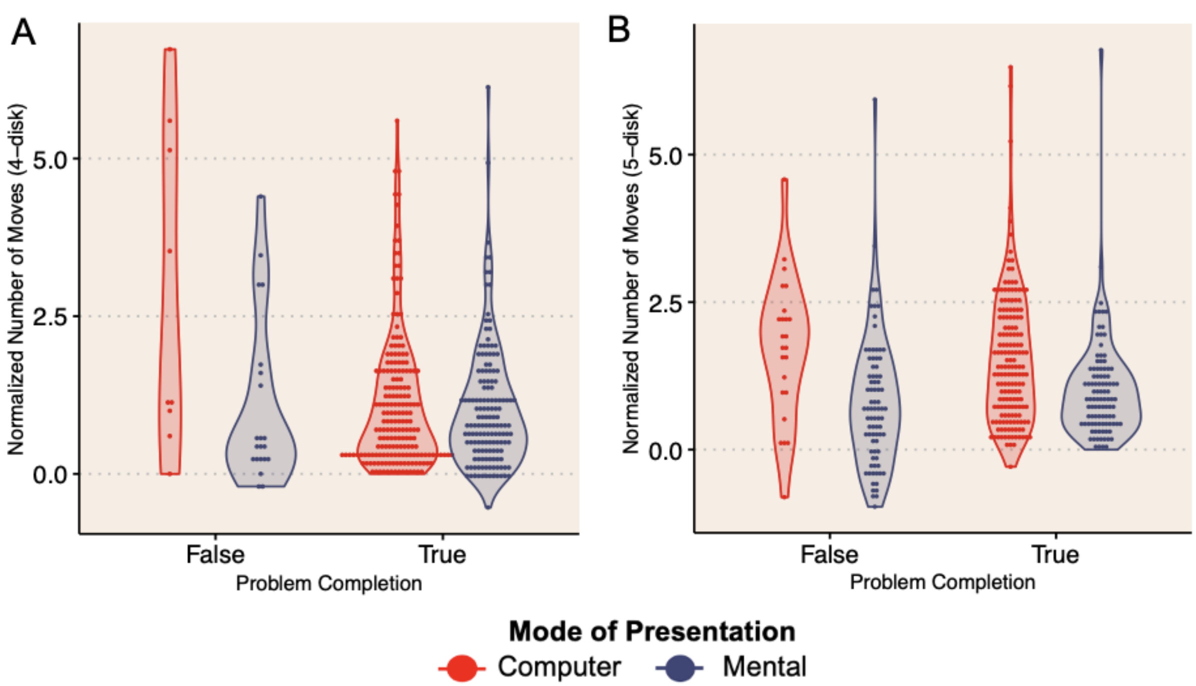

Figure 6. *Illustrates number of moves for those who did (right) and did not (left) complete the 4-disk (A) and 5-disk (B) TOH problems in the allotted 15 min.*

Performance on the second problem (5-disk)

Number of Reminders. The number of reminders participants needed when solving the second problem mentally did not differ between MM (M = 13.34) and CM (M = 16.07) groups (t(178.55) = 1.33, $CI_{95\%}$ = [-1.31,6.78], p = .18). It's worth noting that the number of reminders for these conditions in Experiment 2 is much larger than was reported in Experiment 1 (MM M = 1.21; CM M = 2.18), which may be because participants did not have to make a request of the experimenter, which may have been more intimidating.

Completion rate. As hypothesized, there was a significant main effect of 5-disk mode of presentation such that those solving the 5-disk problem on the computer were more likely to finish the problem (87.37%) than those solving the problem mentally (55.25%) problem (B = -1.75, $SE(B)$ = .33, p < .001). Interestingly, there was also a significant main effect of mode of presentation of the first, 4-disk, problem such that those who solved the 4-disk problem mentally were more likely to complete the 5-disk problem (79.29%) than those who solved the first, 4-disk problem on the computer (65.35%), regardless of mode of presentation (B = .98, $SE(B)$ = .50, p = .0488). There was no significant interaction between first and second problem mode of presentation (B = -.15, $SE(B)$ = .58, p = .79) (see Fig 4D).

Progress to the goal state. There was a main effect of 5-disk mode of presentation such that those in the mental conditions (M = 1.49) made less progress towards to goal state than those in computer conditions (M = 3.63) (B = -2.84, $SE(B)$ = .47, t(367) = -6.03, p < .001). There was no significant main effect of mode of presentation of the first, 4-disk, problem (B = .02, $SE(B)$ = .49, t(367) = .04, p = .97). There was a significant interaction (B = 1.48, $SE(B)$ = .70, t(367) = 2.12, p = .03) between first and second problem mode of presentation, such that those solved the first problem mentally (M = 2.28) made more progress towards the goal state when solving a second problem mentally compared to participants who first solved a problem on the computer (*M* = .78).

Number of moves. Participants who did not complete the 5-disk problem were filtered from this analysis. There was no main effect of mode of presentation of the first, 4-disk, problem (*B* = -.05, *SE(B)* = .16, *t*(262) = -.30, *p* = .76). However, as hypothesized, participants solving the 5-disk problem mentally completed the problem in fewer moves than participants solving the problem on the computer (*B* = -.44, *SE(B)* = .19, *t*(262) = -2.33, *p* = .02). There was no interaction between modes of presentation on the 4 and 5-disk problems (*B* = -.11, *SE(B)* = .26, *t*(262) = -.43, *p* = .67; see Figure 6B for a distribution of the data and 4E for number of move analysis).

Average time between moves. Even though participants did not verbalize their moves in the mental condition of this experiment, participants in the mental condition (*M* = 13.99 sec) still spent more time between moves on the 5-disk problem (*M* = 5.24 sec) (*B* = 9.60, *SE(B)* = 1.33, *t*(367) = 7.24, $p < .001$) relative to the external condition. There was no effect of mode of presentation of the first, 4-disk problem (*B* = .20, *SE(B)* = 1.37, *t*(367) = .15, *p* = .89), nor an interaction between modes of presentation on the 4 and 5-disk problems (*B* = -1.79, *SE(B)* = 1.96, *t*(367) = -.91, *p* = .36; see Figure 4F).

Discussion

Experiment 2 replicated our findings from Experiment 1, namely, that participants were more likely to complete problems when solving them with the computer but spent more time between moves when solving problems mentally, regardless of difficulty. Although the number of reminders increased in the mental condition, the amount of time spent looking at the current state of the problem after pressing the “z” key was removed from the total time before calculating average time per move. It is also worth noting that although participants asked for more reminders in Experiment 2, findings from Experiment 1 were still replicated, suggesting that a problem does not need to be purely “mental” for participants to implement such strategies. We also found that participants solved the more difficult, 5-disk, problem in fewer moves when solving the problems mentally.

These results together suggest that participants solving difficult TOH problems mentally may have used a more sophisticated strategy than those solving the problems on the computer. Lastly, though the effect sizes were small, there was evidence of transfer such that solving an initial problem mentally enhanced performance on the second problem. Specifically, individuals who first completed the 4-disk problem mentally were more likely to finish the 5-disk problem regardless of mode of presentation. Additionally, people who completed the first problem mentally made more progress to the goal state when given a more difficult mental problem compared to people who completed the first problem using an external representation.

General Discussion

Across two experiments, the current work examined the roles of internal and external representation and problem difficulty on solving recursion problems. We hypothesized that participants would be more likely to successfully solve problems when given an external representation but would show evidence of complex strategy use when solving problems mentally.

Experiments 1 and 2 both suggest that external representations help with completing problems as participants were more likely to complete the 4- and 5-disk problems when solving the problem with an external representation across both experiments. The external representation acts as an extension of working memory, creating a task that is less cognitively demanding and therefore increases the likelihood that participants can complete the problem when given a time limit. However, participants in mental conditions completed difficult problems in fewer moves and spent more time between moves, suggesting that they were using a more sophisticated strategy than those in the computer conditions. While participants in the computer condition may have had a different, simpler strategy, that allowed them to reach the goal state, participants were not necessarily “successful” as they completed the problem in more moves than those in the mental condition.

Time Between Moves and Strategy Use

Experiments 1 and 2 suggest that internal representation-based problem-solving may promote the development of more complex strategies. We argued that participants engaged in more complex strategies because they spent more time between moves than participants solving problems on the computer, regardless of difficulty (as in Noyes & Garland, 2003). For example, Welsh et al. (1995) found that optimal problem solvers critically spent more time before making a move at points during the problem where errors are commonly made. We recognize that this measure is an imperfect proxy for complex strategy use. However, it is worth noting that participants in mental conditions spent sometimes twice as much time between moves when compared to participants in computer conditions, even after correcting for articulation in Experiment 1, and, importantly, when they used the identical keyboard entry procedure to communicate moves in Experiment 2.

Number of Moves and Strategy Use

In addition to time-between-moves, number of moves was also used as a proxy for strategy use, as the goal of TOH is to solve the problem in as few moves as possible, and this was explicitly explained to participants. One of the benefits of studying problem solving with

TOH is that there is a reasonably simple optimal solution, and thus, experimental performance may be compared to the optimal number of moves. Across both experiments, participants in mental conditions solved the 5-disk problem using fewer moves than those in the computer conditions. However, this analysis is limited by the fact that we only included participants who solved the problem within the allotted time period, and participants were less likely to solve the 5-disk problem in the mental condition. This discrepancy led to unequal sample sizes in the computer and mental conditions for these analyses, which we attempted to correct for by using regression methods (Slinker & Glantz, 1988). One may argue that only including those who finished the problem created a selection effect in which only the best participants in the mental conditions were compared to all of the participants in the computer condition. Indeed, this is a valid concern. However, as suggested by the distributions of data in Figures 3B and 6B, participants who did not complete the problem in the 15 min period were not unsuccessful due to mindlessly moving the disks, these data are positively skewed with most of these participants being around the optimal number of moves (i.e., normalized score of 0), or less.

There is a speed/accuracy tradeoff when solving problems like TOH (Rönnlund et al., 2001, 2007). These data suggest that different problem representations may bias participants toward speed or accuracy. Participants in the computer conditions spent less time between moves and solved difficult problems using a greater number of moves than those in the mental conditions. It appears that external representations biased participants toward speed, while internal representations biased participants toward accuracy. This bias may have been specific to our research paradigm, as participants in the mental conditions had no choice but to value accuracy due to working memory limitations while participants in computer conditions did not have this limitation. Future work should investigate how increasing motivation in participants influences these results (Eseryel et al., 2014). Although, it is important to remember that in real-life circumstances, it is often the case that people are not sufficiently motivated to value accuracy over speed, for instance, students in a classroom. These data suggest that giving students a problem with less external representation may help with conceptual understanding and strategy generation.

Self-Reported Strategies

The self-report data corroborate the findings from the number of move and time between move analyses. In Experiment 1 participants solving problems mentally reported using *remembering move sequences* more than participants solving problems on the computer. Participants solving problems on the computer reported using *trial and error* more than participants solving problems mentally. These findings were replicated in a follow-up study reported in the Supplemental Materials where participants were randomly assigned to CC, MM, CM, and MC conditions and asked to self-report their strategies after solving a 5-disk problem followed by a 4-disk problem.

Strategy Transfer

These experiments provide promising evidence that first solving a problem high in internal representation leads to better performance on future problems. In both Experiments 1 and 2 participants solving the 4-disk problem mentally made more progress towards the goal state on a second, more difficult problem, when compared to participants who solved the first problem on the computer. In Experiment 1, participants who first solved a 4-disk problem on the computer later spent more time between moves when solving a problem high in internal representation when compared to participants who first solved a problem mentally. Participants solving the 5-disk problem mentally still spent more time between moves than the computer conditions overall, suggesting that participants may be faster at implementing their strategy after solving the first problem mentally. However, it is important to point out that the interaction between first and second problem mode of presentation was not replicated in Experiment 2, suggesting that this specific finding is not robust.

It is possible that the evidence supporting strategy transfer is merely due to transfer appropriate processing (Morris et al., 1977) as it is well documented that problem solving improves if one is assessed in a context similar to the context in which learning occurred (Gick & Holyoak, 1983). One would expect there to be a cost associated with switching from one mode of presentation to another, as strategies that may be effective when solving problems on the computer may be costly when solving problems mentally and vice-versa. However, transfer appropriate processing or a switch cost due to mismatching modes of presentation cannot account for the finding in Experiment 2 that first solving a problem mentally led participants to make more progress towards the goal state on a second problem, regardless of the 5-disk problem's mode of presentation. Future work could examine this possibility further by giving participants a novel test problem in which transfer appropriate processing cannot apply, such as an isomorphic problem (Simon & Hayes, 1976), or a problem with a different mode of presentation (i.e., a physical tower).

Limitations

The main limitation of the current study is the conclusion of complex strategy use based on number of moves to solve the problems, time between moves, and self-report data. Future research should attempt to explicitly study strategy use. For example, one could implement a "think aloud" protocol (Ericsson & Simon, 1980) to examine the strategies participants use or compare participant data to computational models of different TOH strategies (e.g., Anderson et al., 2005; Goel et al., 2001). In addition, when participants were asked to self-report strategy use, these strategies were not defined beyond their labels, unless a participant asked the experimenter for further clarification. In future work, these strategies should be better defined for participants so it is assured that they understand the meaning of each of the strategies listed.

Interpretations of the data are also limited by the fact that number of moves were only examined for those who successfully completed the problem in 15-min. Future work should attempt to measure number of moves in a way that allows all participant data to be considered (i.e., using machine learning algorithms to predict the sequence of moves participants would have used to complete the problem). We also acknowledge that our measure of progress towards the goal state may not be the optimal way to measure this construct. One alternative method may be to consider the minimum number of moves a participant has left until reaching the goal state. The current study did not implement this measure as we did not want to assume that participants would move towards the goal state with an optimal strategy, thus not reflecting the actual number of moves participants had left until completion.

Conclusions

These findings add to the existing literature supporting the idea that internal representations promote complex strategy learning while external representations may promote the use of trial-and-error problem solving and facilitate problem completion (Noyes & Garland, 2003). The current research is also consistent with the idea that sometimes the “effortful struggle” of tasks requiring higher level cognition leads to better conceptual understanding and performance on other tasks (Jensen et al., 2014; Jonsson et al., 2016). It also supports the body of literature suggesting that “desirable difficulties” are ideal for learning, whether that be in the actual difficulty of the problem relative to other problems or in amount of internal or external representation in the problem space (Bjork, 1994).

The current study indicated that participants solving problems high in internal representation formed strategies resulting in fewer moves. However, it did not determine whether those strategies are lasting. Research should use longitudinal methods to see if internal representations promote lasting strategies that lead to more efficient performance when presented with a recursion problem in the future. Internal representations could also aid with forming strategies for problems that may not necessarily require recursion. Future research should examine the role of internal representations in other contexts, such as deductive reasoning, where a similar sub-goal process is required. In addition, it’s important to remember that problem spaces are distributed across internal and external representations. For example, in the experiments reported here, problems solved mentally were more internally represented than problems solved with the computer, however, participants solving the problems internally may have also relied on some external representation as well (i.e., verbalization, gesturing, reminders of the problem state). Future research should examine the external representations recruited when participants are tasked with cognitively demanding, internally represented tasks. Finally, the study of problem solving, higher order cognition, and the role of problem representation should be explored further, and in contexts beyond abstract problems such as TOH, as best said by Karat (1982, pp. 555-556):

> “It is hoped that the success of this effort will aid the study of problem solving and learning in more complex areas. After all, the study of Tower of Hanoi problems is not the goal of cognitive science; however, the advantages of problem-solving study utilizing problems of little practical value, but significant structural advantages, should not be

> overlooked in this effort. The advantages to a problem solver in making the correct first move can be enormous”.

Acknowledgements

The authors thank Dr. Thomas Redick for helping with project design, Tyler J. Adkins for assisting with data pre-processing, and Abbey Roelofs for programming the task used in Experiment 2. This work was supported, in part, by the National Science Foundation Grants #1349677-BCS and #1662230 to SH and a Rackham Graduate Student Research Grant awarded to MF.

References

Adams, H. B. (1986). The teaching of general problem solving strategies. *Gifted Education International*, *4*(2), 85–88. https://doi.org/10.1177/026142948600400205

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*(1), 35–44. https://doi.org/10.1037/0003-066X.48.1.35

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced Tower of Hanoi. *Journal of cognitive neuroscience*, *17*(8), 1261-1274.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612–637. https://doi.org/10.1037/0033-2909.128.4.612

Barrett, T. J., Stull, A. T., Hsu, T. M., & Hegarty, M. (2015). Constrained interactivity for relating multiple representations in science: When virtual is better than real. *Computers & Education*, *81*, 69–81. https://doi.org/10.1016/j.compedu.2014.09.009

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(1), 153–166. https://doi.org/10.1037/0278-7393.15.1.153

Bassok, M., & Novick, L. R. (2012). Problem solving. In *The Oxford handbook of thinking and reasoning*. (pp. 413–432). Oxford University Press.

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185-205). Cambridge, MA: MIT Press.

Bocanegra, B. R., Poletiek, F. H., Ftitache, B., & Clark, A. (2019). Intelligent problem-solvers externalize cognitive operations. *Nature Human Behaviour*, *3*(2), 136–142. https://doi.org/10.1038/s41562-018-0509-y

Cary, M., & Carlson, R. A. (2001). Distributing working memory resources during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(3), 836–848. https://doi.org/10.1037/0278-7393.27.3.836

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1147–1156. https://doi.org/10.1037/0278-7393.15.6.1147

Davis, H. P., & Klebe, K. J. (2001). A longitudinal study of the performance of the elderly and young on the tower of hanoi puzzle and rey recall. *Brain and Cognition*, *46*(1–2), 95–99. https://doi.org/10.1006/brcg.2000.1269

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251. https://doi.org/10.1037/0033-295X.87.3.215

Eseryel, D., Law, V., Ifenthaler, D., Ge, X., & Miller, R. (2014). An investigation of the interrelationships between motivation, engagement, and complex problem solving in game-based learning. *Educational technology & society*, *17*(1), 42-53.

Ford, G. (1984). An implementation-independent approach to teaching recursion. *Proceedings of the Fifteenth SIGCSE Technical Symposium on Computer Science Education - SIGSCE ’84*, 213–216. https://doi.org/10.1145/800039.808655

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist*, *21*(1–2), 99–120. https://doi.org/10.1080/00461520.1986.9653026

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in "planning" functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia*, *33*(5), 623–642. https://doi.org/10.1016/0028-3932(95)90866-P

Goela, V., Pullara, D., & Grafman, J. (2001). A computational model of frontal lobe dysfunction: Working memory and the Tower of Hanoi task. *Cognitive Science*, *25*(2), 287-313.

Greeno, J. G. (1978). Natures of problem-solving abilities. In *Handbook of learning & cognitive processes: V. Human information.* (pp. 239–270). Lawrence Erlbaum.

Handley, S. J., Capon, A., Copp, C., & Harper, C. (2002). Conditional reasoning and the Tower of Hanoi: The role of spatial and verbal working memory. *British Journal of Psychology*, *93*(4), 501–518. https://doi.org/10.1348/000712602761381376

Hayes, J.R., Simon, H.A. (1977). Psychological differences among problem isomorphs. In Castellan, N.J., Pisoni, D.B., Potts, G.R., (Eds.), Cognitive theory (Vol. 2. pp, 21–44), Hillsdale, NJ: Lawrence Erlbaum.

Hélie, S. & Pizlo, Z. (in press). When is psychology research useful in artificial intelligence? A case for reducing computational complexity in problem solving. *topiCS in Cognitive Science*.

Jacewicz, E., Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language variation and change*, *21*(2), 233.

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test…or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, *26*(2), 307–329. https://doi.org/10.1007/s10648-013-9248-9

Jonsson, B., Kulaksiz, Y. C., & Lithner, J. (2016). Creative and algorithmic mathematical reasoning: Effects of transfer-appropriate processing and effortful struggle. *International Journal of Mathematical Education in Science and Technology*, *47*(8), 1206–1225. https://doi.org/10.1080/0020739X.2016.1192232

Karat, J. (1982). A model of problem solving with incomplete constraint knowledge. *Cognitive Psychology*, *14*(4), 538–559. https://doi.org/10.1016/0010-0285(82)90018-4

Klahr, D., & Chen, Z. (2011). Finding one's place in transfer space. *Child Development Perspectives*, *5*(3), 196–204. https://doi.org/10.1111/j.1750-8606.2011.00171.x

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology*, *13*(1), 113–148. https://doi.org/10.1016/0010-0285(81)90006-2

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*(2), 248–294. https://doi.org/10.1016/0010-0285(85)90009-X

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

Lucas, E., & Claus, N. (1883). The Tower of Hanoi. Paris: GAUTHER-VILLARS, printer of the Academie des Sciences and the Ecole Polytechnique Quai des Augustins, 55.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

Noyes, J. M., & Garland, K. J. (2003). Solving the Tower of Hanoi: Does mode of presentation matter? *Computers in Human Behavior*, *19*(5), 579–592. https://doi.org/10.1016/S0747-5632(03)00002-5

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *39*(2), 240–272. https://doi.org/10.1037/h0080061

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, *6*(3), 436–450. https://doi.org/10.1016/0010-0285(74)90020-6

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning sciences*, *13*(3), 273-304.

Reuter-Lorenz, P. A., Jonides, J., Smith, E. E., Hartley, A., Miller, A., Marshuetz, C., & Koeppe, R. A. (2000). Age differences in the frontal lateralization of verbal and spatial working memory revealed by PET. *Journal of Cognitive Neuroscience*, *12*(1), 174-187.

Rönnlund, M., Lövdén, M., & Nilsson, L. G. (2001). Adult age differences in Tower of Hanoi performance: Influence from demographic and cognitive variables. *Aging, Neuropsychology, and Cognition*, *8*(4), 269-283.

Rönnlund, M., Lövdén, M., & Nilsson, L. G. (2007). Cross-sectional versus longitudinal age gradients of Tower of Hanoi performance: The role of practice effects and cohort differences in education. *Aging, Neuropsychology, and Cognition*, *15*(1), 40-67.

Rowe, G., Hasher, L., & Turcotte, J. (2008). Age differences in visuospatial working memory. *Psychology and Aging*, *23*(1), 79–84. https://doi.org/10.1037/0882-7974.23.1.79

Schiff, R., & Vakil, E. (2015). Age differences in cognitive skill learning, retention and transfer: The case of the Tower of Hanoi Puzzle. *Learning and Individual Differences*, *39*, 164–171. https://doi.org/10.1016/j.lindif.2015.03.010

Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. *Informal reasoning and education*, 311.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, *7*(2), 268–288. https://doi.org/10.1016/0010-0285(75)90012-2

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitivepsychology*, *8*(2), 165-190.

Slinker, B. K., & Glantz, S. A. (1988). Multiple linear regression is a useful alternative to traditional analyses of variance. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *255*(3), R353-R367.

Spencer, R. M., & Weisberg, R. W. (1986). Context-dependent effects on analogical transfer. *Memory & Cognition*, *14*(5), 442–449. https://doi.org/10.3758/BF03197019

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Welsh, M., Cicerello, A., Cuneo, K., & Brennan, M. (1995). Error and temporal patterns in Tower of Hanoi performance: Cognitive mechanisms and individual differences. *The Journal of General Psychology*, *122*(1), 69-81.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, *17*(2), 89-100.

Wu, H.K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, *88*(3), 465–492. https://doi.org/10.1002/sce.10126

Zhang, J, & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122. https://doi.org/10.1207/s15516709cog1801_3

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217. https://doi.org/10.1207/s15516709cog2102_3

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition*, *56*(3), 286–292. https://doi.org/10.1016/j.bandc.2004.07.003

Appendix

1. The problems I solved today were difficult (7-pt Likert Scale; Strongly Disagree – Strongly Agree)

2. The problems were more difficult when the number of disks increased (7-pt Likert Scale; Strongly Disagree – Strongly Agree)

3. Please report the strategies that you used when solving the problems and the percent of time you used each strategy:
    a. Remembering move sequences (0-100 slider scale)
    b. Developing step-by-step instructions to solve the problem (0-100 slider scale)
    c. Use of sub-goals (0-100 slider scale)
    d. Trial and error (0-100 slider scale)

4. I believe that I could apply my strategy to the problem given any number of rings (7-pt Likert Scale; Strongly Disagree – Strongly Agree)

5. To what degree do you believe that you understand the problem structure? (7-pt Likert Scale; Not at all – Completely)