AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

2022, Vol. 48, No. 3, 348–374
https://doi.org/10.1037/xlm0001117

The Role of Domain-General Attention and Domain-Specific Processing in Working Memory in Algebraic Performance: An Experimental Approach

Zehra Emine Ünal, Alicia Forsberg, David C. Geary, and Nelson Cowan
Department of Psychological Sciences, University of Missouri, Columbia

We investigated the role of working memory in symbolic and spatial algebra and related tasks across five experiments. Each experiment combined a processing task (expression evaluation, arithmetic, coordinate plane, geometry, or mental rotation) with verbal and spatial memory loads in a dual-task design. Spatial memory was compromised in the presence of more difficult processing tasks, and verbal memory was only compromised in the presence of algebraic tasks. The latter was related to the demands of retaining quantities associated with variables in verbal memory. We suggest that both verbal and spatial working memory retention engage domain-general attention, but that their maintenance mechanisms differ. Verbal memory has attention-based and rehearsal-based mechanisms, and thus sustaining verbal information over a short period is less attention-demanding than holding spatial information. We suggest that effects of a memory load on processing (e.g., $x = 6$) depend on whether use of maintenance strategies are possible for the specific memory load while carrying out processing. In all, our results indicate that algebraic tasks use domain-general attention and include verbal processing of algebraic variables (i.e., information conveyed in x, y). We discuss the implications for algebra learning and working memory theories.

Keywords: algebra, working memory, domain-general attention, verbal processing

Competence with algebra sets the foundation for learning the more complex mathematics that undergirds preparation for success in many science, technology, engineering, and mathematics (STEM) fields (Gamoran & Hannigan, 2000; National Mathematics Advisory Panel [NMAP], 2008) and contributes to future career success more broadly (Crisp et al., 2009; Hansen, 2014). Despite its well-documented importance and repeated attempts to improve outcomes, many students fail to achieve a level of competence with algebra that prepares them for further mathematics learning (Stein et al., 2011). Changes in instructional approaches are needed for these students and basic research on the factors that contribute to algebraic learning has the potential to influence the development of such approaches, but there are few such studies (e.g., Geary et al., 2015; Pollack et al., 2016; Sweller & Cooper, 1985; Walczyk & Griffith-Ross, 2006).

Based on cognitive load theory, instructional practices that reduce the working memory demands of the presented material foster learning across domains, especially during the early phases of this learning (Cooper & Sweller, 1987; Paas et al., 2010; Sweller et al., 2019; Van Merrienboer & Sweller, 2005). The development of such instruction approaches will be facilitated by an understanding of the working memory and attentional demands associated with performance in key aspects of algebra. Across five experiments, we explore the working memory and attentional demands associated with early but critical components of algebraic knowledge and skills; specifically, efficiency and accuracy in evaluating expressions and placing ordered pairs in the coordinate plane, both of which predict performance on tests of algebra achievement (Geary et al., 2015).

The Current Study

Although numerous studies have demonstrated a significant relationship between working memory and mathematics and linked higher working memory capacity with better mathematics performance (Peng et al., 2016), most of the associated findings have come from individual differences research, and thus, causal inferences cannot be drawn. Moreover, despite the considerable importance of advanced mathematics skills, including algebra, less is known about learning in these areas relative to arithmetic (Peng

Alicia Forsberg https://orcid.org/0000-0002-6635-6160
David C. Geary https://orcid.org/0000-0003-3029-6343
Alicia Forsberg is now at the Department of Psychological Sciences, University of Sheffield.

This project was supported by DRL-1659133 from the National Science Foundation and R01 HD087231 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (National Institute of Child Health and Human Development) to David C. Geary, and National Institute of Child Health and Human Development Grant R01-HD021338 to Nelson Cowan. We thank Lutfiye Tutuncu, Busra Unal, Hafize Unal, Gamze Kartal, Brett Glass, Mary Hoard, Lara Nugent, Dominic Guitard, Tuba Isbilir, Selin Sancaktar, Ayse Turkucar, Eryn Adams, Kristin Balentine, and Joe LaMendola for their help during experiment preparation.

Data and analyses can be found at https://osf.io/g6mak/ (Ünal, 2021).

Correspondence concerning this article should be addressed to Zehra Emine Ünal, Department of Psychological Sciences, University of Missouri, Columbia, 317 McAlester Hall, Columbia, MO 65211-2500, United States. Email: zeuqq6@umsystem.edu

et al., 2016). Almost none of the studies that have been conducted, to our knowledge, has used experimental designs to identify the working memory processes engaged during problem solving in these areas (though effects of memory load on algebraic processing were observed by Ellis et al., 2020).

We aim to address this gap by investigating the working memory processes in higher-level mathematics through experimental manipulations, using a dual-task method. The latter requires performing a primary task (i.e., processing task) under a memory load. Our focus is specifically on algebra because of its foundational role in more complex mathematics (Fyfe et al., 2018; Gamoran & Hannigan, 2000; NMAP, 2008) and the coordinate plane as a precursor of algebra (Leinhardt et al., 1990).

Current Processing Tasks

We conducted five experiments, each with different processing tasks (i.e., expression evaluation, arithmetic, coordinate plane, geometry, and mental rotation) and two memory loads (i.e., verbal and spatial). The arithmetic task was included to assess the working memory and attentional demands of the arithmetic in the expression evaluation task and the geometry and mental rotation tasks were included to assess the potential spatial load demands of the coordinate plane task. As an experimental procedure, we first present a memory load (e.g., word list), then ask participants to do a processing task and, after that, asked them to recall the memory load material.

Current Memory Load Tasks

We chose commonly used, simple, verbal, and visual short-term memory (STM) tasks to serve as memory loads during processing. In the verbal memory task, participants saw a sequence of six vertically represented words (i.e., reference array). The words disappeared and then reappeared after 16 s. The participants were to determine whether the final word sequence was in the same order as the initially presented sequence (for a similar procedure, see Trbovich & LeFevre, 2003). In the spatial memory task, participants saw six crosses in a four-by-four grid, followed by an 11.3-s blank screen. They then determined whether a newly presented group of six crosses was in the same or different position as the original group. The type of visual load was like the one used in the visual pattern span task (Della Sala et al., 1999).

These tasks were meant to be relatively comparable except that the verbal task involves word sequences, and the visual task involves a spatial array of marked locations. One can compare the tasks on both the types of materials involved and the types of processes involved. One point of view is that verbal and visual materials are stored in different modules (e.g., Baddeley & Hitch, 1974; Baddeley & Logie, 1999), and thus the different types of memory load should differentially affect processes that draw upon verbal and visual stores. For example, mental rotation must involve spatial processing that should be dependent on a visual store, whereas the assessment of algebraic expressions must involve symbolic processing that should be dependent on a verbal store.

In addition to these modality effects, however, an asymmetry would be expected according to other points of view (e.g., Cowan, 1988, 2019), in which attention plays an important role. Specifically, there is considerable research support for an asymmetry in which the storage of visual materials recruits general attention, whereas verbal materials recruit it less and can make use of less-attention-demanding, verbal forms of holding information (Barrouillet & Camos, 2015; Gray et al., 2017; Morey et al., 2013; Vergauwe et al., 2010). According to these views, there should be a more pervasive effect of visual span than of verbal span. The two span tasks used here were selected to assess the processes involved in each kind of algebraic task, including both domain-specific processes and a reliance on general attention.

A great deal of previous research confirms that the effect of processing tasks on memory load retrieval depends on task demands. For instance, random interval generation (i.e., production of time intervals that do not follow any pattern; Vandierendonck et al., 1998) interfered with verbal and spatial span tasks (Martein et al., 1999). On the other hand, articulatory suppression (i.e., repetition of the same verbal stimuli, DeStefano & LeFevre, 2004) selectively affected memory span task performances; participants could recall fewer verbal items, but it didn't influence the maximum number of spatial items they remembered (Martein et al., 1999). The observation of interference between a processing task and a memory load task would indicate shared resources (Trbovich & LeFevre, 2003) if both tasks demand general attention. Therefore, using the same memory load tasks from one experiment to the next, a difference in the extent of mutual interference between a visual load and processing may indicate a difference in processing task difficulty. Any effect from a dual-task cost between visual memory and the processing task also could be interpreted as indicating a specific visual processing component of the task, but that conclusion is more likely for some processing tasks than for others. For example, interference between a visual load and mental rotation could easily be due to shared visual processing, but the same cannot be said if there is a dual-task cost between visual span and algebraic expression evaluation. Therefore, our conclusions depend on comparison of the results from all the experiments.

Main Issues

Using this technique, we explored (a) the extent to which coordinate plane and algebra depend on domain-specific stores like the phonological loop and the visual-spatial sketchpad (or in Cowan's approach, activated long-term memory with feature-specific interference) and (b) to the extent to which coordinate plane and algebra depend instead on attention and executive functions. The difference between the two possibilities is that only the latter predicts interference between a verbal process and a visual memory load, or between a visual process and a verbal memory load. Thus, the experiments contribute to our understanding of the working memory and attentional demands of core algebraic competencies and to our understanding of whether these working memory and attentional demands are domain-specific or based on domain-general attentional resources. First, we examine what the literature tells us about what we might expect.

Working Memory and Processing Mathematics Information, Including Algebra

Both experimental and individual differences studies have shown that higher working memory (WM) capacity is associated with better performance in various areas of mathematics, although most of these studies have focused on arithmetic (Adams & Hitch, 1997; Cragg et al., 2017; De Smedt et al., 2009; Geary et al., 2007; Geary & Widaman, 1992; Hitch, 1978; Imbo & Vandierendonck, 2007; for a review, see DeStefano & LeFevre, 2004). Individual differences studies are useful for identifying relations between WM and performance in mathematics, but experimental manipulations of working loads are needed to make causal inferences (e.g., Imbo & Vandierendonck, 2007). Manipulations of visuospatial and verbal loads using dual task methods have been shown to influence aspects of arithmetic computation (DeStefano & LeFevre, 2004; Fürst & Hitch, 2000; Hitch, 1978). As an example, Trbovich and LeFevre (2003) found that manipulation of verbal load disrupted adults' ability to correctly solve horizontally presented arithmetic problems (e.g., 6 + 37), whereas manipulation of spatial load disrupted their ability to solve vertically presented problems.

However, there are no comparable studies (to our knowledge) of visuospatial and verbal load manipulations as related to performance on algebraic tasks. We provide these studies and focus on algebraic expression evaluation and accuracy in placing ordered pairs in the coordinate plane. The latter is an important prerequisite for understanding how equations and functions map to coordinate space. The evaluation of expressions is dependent on the processing of arithmetic and variables (Walczyk & Griffith-Ross, 2006). We first review the evidence on the role of working memory in algebraic expressions and arithmetic evaluations, and then review the evidence on its role for spatial aspects of algebra.

Algebraic Expression and Arithmetic Evaluation

An understanding of and fluency with variables are cornerstones of the early learning of algebra and taught based on students' knowledge of arithmetic (e.g., Barbieri & Booth, 2020; Fuchs et al., 2012; NMAP, 2008; Pillay et al., 1998). Fuchs et al. (2012) showed that students' emerging understanding of variables is influenced by their earlier competence with arithmetic, with individual differences in verbal working memory indirectly influencing the learning of variables through individual differences in arithmetic competence. Walczyk and Griffith-Ross (2006) found that adults' accuracy in solving algebraic expressions (e.g., $x = 6$, solve 3×-6) was strongly correlated ($r = .68$) with the speed of solving complex arithmetic problems (e.g., 82 + 34). Their hypothesis was that arithmetic fluency reduced the WM demands of evaluating the expressions. We might then expect that the influence of WM on evaluating expressions is like the influence of WM on learning and solving complex arithmetic problems.

However, the interference between the verbal load and the processing of arithmetic versus algebraic expressions could differ inasmuch as some recent studies indicate differences in how this information is processed. As an example, Pollack et al. (2016) used a comparison task with Arabic numerals (e.g., 6) and variables (e.g., x, y, and z). After learning and practicing predetermined variable and numeral pairs (e.g., $y = 9$), participants were primed with an Arabic numeral (e.g., 2) and asked to determine whether it was larger than 5. The sequence of priming-target symbols had two different versions: (a) Arabic numeral-Arabic numeral (e.g., 2 then 4), and (b) Arabic numeral-variable (e.g., 7 then W). The goal was to compare the distance effect, in which reaction times are slower and less accurate as the priming and target items get closer to a benchmark (5 in this case), as well as the priming distances effect in which reaction times and error rates go up as the distance between prime and target decreases. It was found, though, that both distance effects were present for the comparison of two Arabic numerals to 5, but absent when the target item was a variable. Based on this finding, the researchers suggested that the mental processing of variables differs from that of Arabic numerals. These results indicate that interpretation of the variables (e.g., retrieving the fact that $G = 7$) may include additional verbal processing and partially overlap in time with the process of the numeral comparisons that drive the distance effect. One possibility is that the overlap obscures the distance effects.

Similar distinctions between numeral and variable processing were observed in magnitude comparisons (Pollack & Price, 2020). While comparing an Arabic numeral (e.g., 5) with a variable (e.g., R, where participants previously learned that $R = 7$), a number word (e.g., TWO), or an artificial symbol (e.g., ▷|, previously learned that ▷|=1), the fastest response times occurred for number words and the slowest for variables. Getting the fastest responses in Arabic numeral-number-word pair comparisons was expected based on participants' long-term exposure to numerals and number words and the quantities they represent. However, slower responses with variables compared with the artificial symbols is noteworthy because it indicates a potential interference effect from literacy experiences when processing variables in a mathematical context (Pollack & Price, 2020; see also McNeil et al., 2010).

Taken together, these findings suggest verbal processing might be uniquely engaged during the solving of algebraic versus arithmetic expressions, given that only the former includes variables. If so, then imposition of a verbal WM load should disrupt the processing of algebraic expressions more severely than the processing of arithmetic expressions.

Spatial Aspects of Algebraic Processing

Competence with the coordinate plane is a significant precursor of and contributor to algebraic learning (Geary et al., 2015), and it is, at the same time, an important part of the transition from number lines to functional graphs (Earnest, 2015). In the initial stages of mathematics development, students form space and number connections through the number line in one-dimensional space where every number corresponds to a point (Earnest, 2015). Later, around fifth grade, they expand this association to the two-dimensional coordinate plane, where number-pairs (i.e., ordered x and y pairs, respectively) describe the position of a point in space (Earnest, 2015). Determining the position of a point on the coordinate plane is vital because it is necessary for graphically presenting the relationship between x and y in algebraic equations (e.g., how y changes while x changes) and strengthens the associations between number and algebra (Kilpatrick et al., 2001; Leinhardt et al., 1990). It has been demonstrated that accuracy in placing ordered pairs on the coordinate plane predicts performance in algebra, controlling other factors (e.g., intelligence, executive functions; Geary et al., 2015), but little is known about the WM demands of making these placements.

Verbal WM is one potential mechanism because, as noted, it appears to be important for processing variables and the order-pairs involve associating numerals with position along the x axis and the *y* axis. We also added a control (spatial) experiment that only required the identification of angle change in the coordinate plane (i.e., *x* axis and *y* axis information did not need to be retained) to investigate whether the spatial nature of the task interferes with the verbal WM memory demand of retaining and placing the coordinate pairs.

It is also possible that visuospatial WM contributes to accuracy in placing ordered pairs because of the spatial nature of the plane. However, visuospatial WM manipulations are not as straightforward as verbal WM manipulations, because it is difficult to separate visuospatial WM demands from general attentional demands. On the one hand, some researchers propose that WM consists of distinct modules with their own attentional resources (e.g., Baddeley & Logie, 1999). On the other hand, opponents of multicomponent working memory models claim that different types of codes are part of activated central working memory, so they are in a common place rather than in separate modules (e.g., Cowan, 1999), or that a representation-refreshing process is often used that requires attention and can be applied to all kinds of representations (e.g., Vergauwe et al., 2010). In other words, a visuospatial WM load might undermine speed and accuracy of placing ordered pairs on the coordinate plane, but the underlying mechanisms would be less certain than for a verbal WM disruption of performance.

Placing ordered pairs in the coordinate plane requires, we expect, the focusing of attention on a discrete area of the plane, as well as coordinating information about relative position along the *x* and *y* axis to make an accurate placement; placement of magnitudes on the *x*-axis is related to visuospatial attention (Geary et al., 2021). We were interested in the latter and thus conducted a control experiment that involved a mental rotation task because of its well-documented spatial component, absence of a verbal component, and necessity of intensive attention (Berch et al., 1998). We hypothesize that either storage (i.e., memory loads) or processing task (i.e., mental rotation) can use attention, and thus storage and processing task interfere with one another even when they occur in different modalities, as in Vergauwe et al. (2010).

Distinguishing Between Views of Working Memory

As summarized in the third and fourth columns of Table 1, there is a way to determine the most probable mechanism or mechanisms related to the different processing task demands. We suggest that there are substantial verbal processing demands in the first three experiments (i.e., expression evaluation, arithmetic equations, coordinate plane), and substantial spatial processing demands in the last three experiments (i.e., coordinate plane, geometry, mental rotation). If verbal and spatial representations engage different modules with independent resources (Baddeley & Logie, 1999), then each type of processing task should result in interference with the kind of load or loads that match the processing task demands.

In contrast, if there only is a single attentional mechanism that is engaged in all tasks, we should see interference between storage and processing for all tasks. Intermediate between these views, Cowan (1988, 1999, 2019) has suggested that there is a need for attention in most tasks but also with a second, less-attention-demanding mechanism of verbal rehearsal available when articulation is relevant (see also Barrouillet & Camos, 2015; Gray et al.,

Table 1
The Summary of the Main Experimental Conditions, Task Demands, and Results

Experiment	Task description	Verbal processing requirement	Spatial processing requirement	Verbal load effect on processing?	Spatial load effect on processing?	Process effect on verbal memory load?	Process effect on spatial memory load?
1. Expressions	Decide if the given algebraic expression is correct using previously shown algebraic variable values	Number-word correspondence; Number-algebraic variable match	None	(No)BF = 0.15	(No)BF = 0.12	YESBF = 228.5	YESBF = 19.9
2. Arithmetic	Evaluate if the arithmetic expression is correct	Number-word correspondence	None	NoBF = 0.08	NoBF = 0.05	(No)BF = 0.17	YESBF = 23.8
3. Coordinate Plane	Decide if the location on the coordinate plane is correct using previously shown algebraic variable values (*x* for *x*-axis, *y* for *y*-axis).	Number-algebraic variable match	Ordered pair- location (on the 2D space) match	BF = 0.97[a]	(Coordinate alone > coordinate w/spatial)BF = 6.00	YESBF = 65.0	BF = 1.60[a]
4. Geometry	Decide if the two angles between parallel/unparallel lines are the same	None	Angle-angle match	(Geometry alone < geometry with verbal) BF = 3.25	(No)BF = 0.22	(No)BF = 0.22	(No)BF = 0.11
5. MentalRotation	Decide if the two 3D drawings are the same or not	None	Shape manipulation; Shape-shape match	BF = 1.26[a]	(No)BF = 0.12	NoBF = 0.06	YESBF = 7,307.2

Note. Parentheses indicate a Bayes Factor between 1/10 and 1/3 for the null hypothesis and between 3 and 10 for the alternative hypothesis.
[a] In these cases, a directional indicated that if there is an effect, that it is more likely to be in the expected direction (with the detrimental effect of the load).

2017; Morey et al., 2013; Vergauwe et al., 2010). According to that view, there should always be interference between processing and a visual memory load, but with selective interference between processing and a verbal memory load when the processing requires verbal rehearsal.

There are two different ways in which a dual task can be executed. In an attention-switching method, participants can cease to try to maintain the load while the process is carried out, and then return to the load later, using attention to retrieve information from the activated portion of long-term memory. In contrast, in an attention-sharing method, the participant can attempt to continue to maintain the load while processing. Attention-switching should result in an effect of processing on storage, but little or no effect of storage on processing (cf. Cowan et al., 2021; Doherty et al., 2019). In contrast, the attention-sharing method should result in a more moderate effect of processing on storage (e.g., a more moderate influence of performing a processing task on spatial memory recall) but also an effect of storage on processing (e.g., the influence of holding spatial information on processing task performance; Cowan et al., 2021). Although this distinction could apply to either the domain-specific or the domain-general views of working memory, it is an important one in understanding the full pattern of dual-task effects.

Experiment 1: Expression Evaluation

Method

Participants

We initially recruited 60 participants through Prolific but excluded three of them for their existing or potential distraction statements (e.g., Please try to minimize distractions. If there is a distractor that you cannot help, please tell us what it is; "People talking around me"), reducing the total to 57 participants (38 females, 18 males, one unreported). We simulated data from three imaginary populations, one with a smaller condition effect (H1A), one with a larger condition effect (H1B), and one without a condition effect (H0). For each simulation, we created a large data sample including 1000 imaginary participants with 10 experimental trials in each condition (Conditions A and B). This large data sample had a specified overall accuracy rate for each condition, representing the assumed true population effect for these three different scenarios (H1A, H1B, H0). For each scenario, we selected $N = 52$ random imaginary participants (the smallest included sample size in the article) and ran our statistical model 100 times. This allowed us to assess the rate at which evidence for or against an effect was observed in these 100 simulations.

We consider either a Bayes Factor > 3 and > 10 as evidence for or against an effect. First, in the imaginary population without a condition effect (H0; simulated data sample: condition A: 80% accurate, condition B: 80% accurate), with a Bayes Factor > 3 cut-off, 0% of the simulations found evidence for an effect, 5% of simulations were inconclusive, and 95% found evidence against an effect, suggesting that this sample size and method produced a low rate of false positives. Using a cut-off Bayes Factor > 10, 0% of the simulations found evidence for an effect, 24% of simulations were inconclusive, and 76% found evidence against an effect. Next, we assume a small effect (H1A; condition A: 80% accurate, condition B: 75% accurate). With a cut-off Bayes Factor > 3, 22% of the simulations found evidence for an effect, 23% of simulations were inconclusive, and 55% found evidence against an effect. Using a cut-off Bayes Factor > 10, 13% of the simulations found evidence for an effect, 57% of simulations were inconclusive, and 30% found evidence against an effect. This suggests that if an effect exists but is small (five percentile units or less), we are not very likely to accurately detect it. Finally, we assumed a true larger condition effect (H1B; condition A: 80% accurate, condition B: 70% accurate). With a cut-off Bayes Factor > 3, 78% of the simulations found evidence for an effect, 18% of simulations were inconclusive and 4% found evidence against an effect. Using a cut-off Bayes Factor > 10, 66% of the simulations found evidence for an effect, 33% of simulations were inconclusive, and 1% found evidence against an effect. This suggests that this sample size and method is quite suitable to accurately detect larger effects.

The mean age of participants was 24.03 years ($SD = 3.79$, Range 18.11–31.10). All participants were native English speakers, and most of them (88%) had at least one college degree. Seven percent of the participants were Hispanic or Latino, and 3% did not report their ethnic identity. Their racial distribution was 78% White or European, 2% Asian, 9% Black or African American, and 9% multiracial; 2% did not report their racial identity. To accommodate participants' preferences, all the experiments were compatible with both touchpad (i.e., small screen and tablet) and computer screens (i.e., large screen); 2% and 98% of participants used small and large screens, respectively. They had normal or corrected vision. The study was approved by the University of Missouri Institutional Review Board (IRB #2002634).

Design

The experimental setup is depicted in Figure 1. There were five experimental blocks, three of which consisted of a single task (i.e., either a processing or a memory task), and the other two included a dual task (i.e., a processing task with a memory task). Each block consisted of 10 trials, creating 50 total trials (two blocks of 10 dual-task trials and three blocks of 10 single-task trials).

Single task blocks:

1. Single-task processing comprising 10 expression evaluation questions.

2. Single-task verbal memory comprising 10 verbal memory questions.

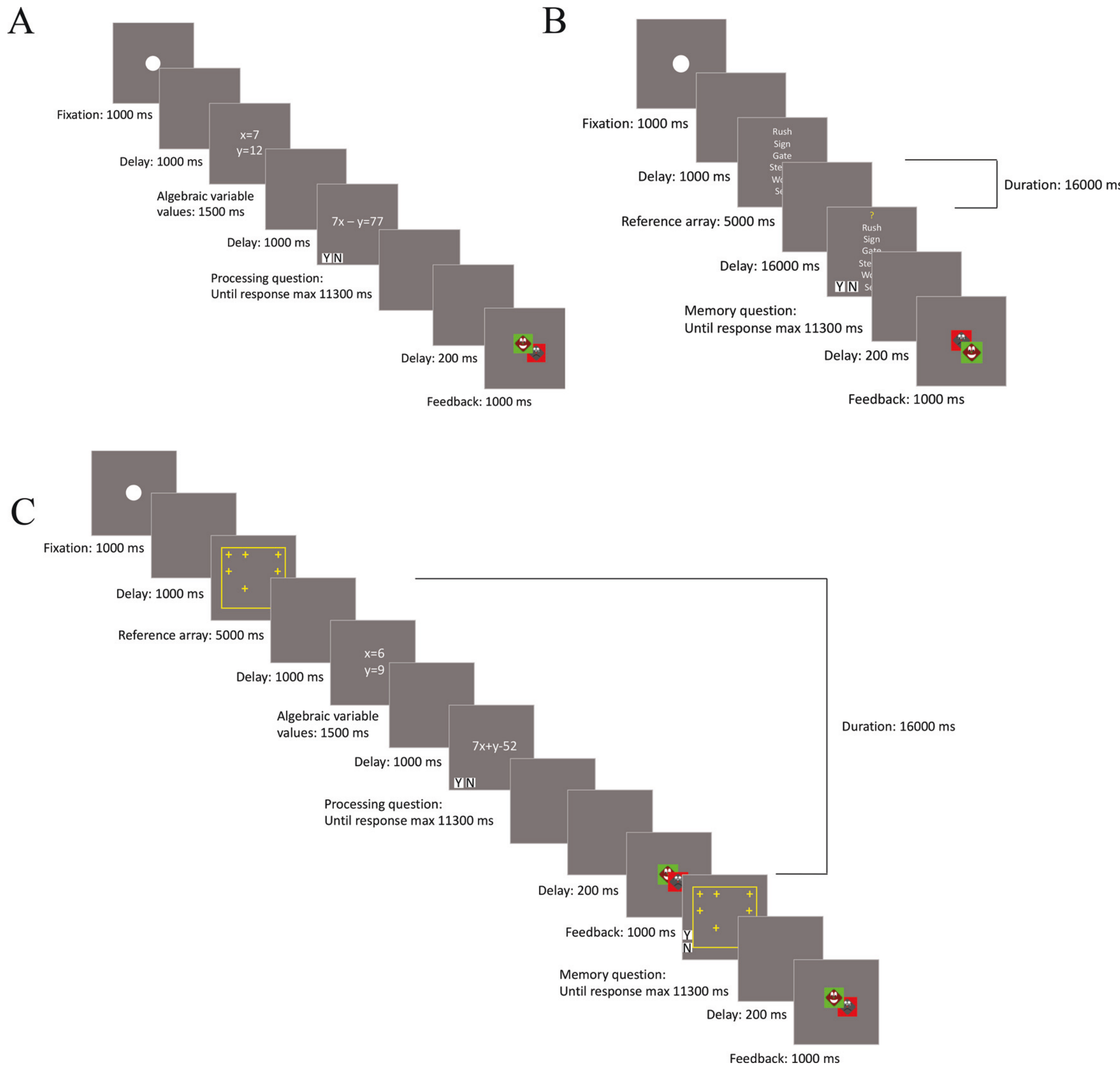
3. Single-task spatial memory comprising 10 spatial memory questions.

Dual tasks blocks:

1. Dual-task verbal memory comprising 10 verbal memory questions and 10 expression evaluation questions. The processing questions were between the reference word array (below) and verbal memory question.

2. Dual-task spatial memory comprising 10 spatial memory questions and 10 expression evaluation questions. The expression evaluation questions were between the reference array (below) and spatial memory questions.

Figure 1
Examples of Three Kinds of Trials in Experiment 1



Note. (A) An example of a single processing task trial (i.e., expression evaluation). (B) An example of a single memory task trial. (C) An example of a dual task trial. See the online article for the color version of this figure.

Within each block, the answer was "yes" (e.g., correct solution) to half of the questions and "no" (incorrect solution) to the other half. Two demonstrations were presented at the beginning of the experiment to introduce the processing and memory tasks. Participants then completed two practice trials to become familiar with the questions and their order within the current block. The order of the blocks and the questions within the blocks was presented randomly for each participant.

Tasks

Expression Evaluation. In the expression evaluation processing task, participants determined (yes, no) whether the solution of the given algebraic equation was correct (or not) for the presented variable values (i.e., expression evaluation, e.g., $7x + y = 52$, given that $x = 6$, $y = 9$).

Immediate Memory Tasks. We chose verbal and spatial STM tasks, with items to be held in memory while doing a processing task, or during a blank interval of the same duration. In the verbal memory task, participants saw a sequence of six vertically represented words (i.e., reference array). The words disappeared and then reappeared after 16 s. Next, the participants determined whether the final word sequence was in the same order as the initially presented sequence.

In the spatial memory task, participants saw six crosses in a four-by-four grid, followed by an 11.3-s blank screen. They then determined whether a newly presented group of six crosses was in the same or different position as the original group. All participants responded by choosing either the Y button for yes or the N button for no, and if they changed their decision, they could deselect their previous choice. The trial ended participants made their final choice (Yes, No).

Materials

We coded all the experiments via the PsyToolkit (Stoet, 2010, 2017). The background color was black, and text colors were white for stimuli and white, yellow, and green for instructions. The text font was Arial 40 pt.

Expression Evaluation. For expression evaluation items, we determined algebraic variable values using one- (e.g., $x = 9$) and two-digit numbers (e.g., $y = 12$) < 20. All expression evaluation items required either carrying or borrowing and included multiplication, as well as either addition or subtraction (e.g., $6y - x = 64$, for the full list of items, see Appendix A). When the equation was solved using the given variables, half of the answers shown matched the solution when properly calculated and half did not. Both correct and incorrect equations always had positive integers as solutions.

Word Lists. For verbal memory tasks, we prepared 10 groups of six words for each block from a list of 120 one-syllable, three- to five-letter words. No word was used more than once. We used the MRC Psycholinguistic Database to generate the word list (Wilson, 1988). The frequency levels of the words were between 551 and 646, meaning that they were among the most frequently used of all words. All the words functioned exclusively as nouns, not as both nouns and verbs (e.g., "walk"). After choosing the words, we randomized the entire sequence of 120 words and divided them into groups of six for 20 trials, using dplyr (Hadley et al., 2019) and tidyr (Hadley & Henry, 2019) packages in R (R Core Team, 2019). We presented all word sequences horizontally.

Spatial List. For the spatial list, we used a 4×4 grid and assigned a number between 1 and 16 to each cell, starting with 1 on the top-left corner and increasing from left to right. We then used dplyr (Hadley et al., 2019) and tidyr (Hadley & Henry, 2019) packages in R (R Core Team, 2019) to choose six of the 16 numbers randomly (e.g., 3, 7, 15, 13, 2, 8). Yellow crosses were then added to the corresponding cells. We repeated this process for all 20 six-cross arrangements. While constructing the patterns, we chose only those that had no more than two adjacent crosses.

Instructions

Each experiment began with general instructions, including information regarding the experimental design (e.g., type of tasks, pictures of feedback, and purpose of the buttons). After that, the participants received task-specific explanations before the demo; The goal was to familiarize participants with the processing and memory tasks. These explanations contained both verbal and pictorial information about the task contents and experimental procedures. Following the demonstration, we informed participants that all tasks in the experiments (dual vs. single) were equally important ("Please remember, during the experiment, you will sometimes get some of these tasks in combination with each other. The other times, you will do a single task. However, all tasks are equally important.") Then, the experimental section began, which included training trials and test sections. Before both training trials and test sections (across all conditions), the notification message appeared on the screen, together with a request to be as fast and accurate as they can ("Please respond as fast and accurate as you can! This task is about speed and accuracy!"). All task instructions are in the additional materials.

Procedure

The procedure and timing of stimuli are shown in Figure 1. The duration between the disappearance of the reference memory array and the appearance of the test memory array was 16 s for both single-task memory blocks and dual-task blocks. The time intervals were equal across conditions to determine if inserting a processing task between collections to-be-remembered (e.g., word list) and memory questions affected recall of the memory items. In the dual-task blocks, the maximum amount of time used to complete the processing tasks was 11.3 s after subtracting the time necessary to present stimuli. As we wanted to keep the time interval constant, participants had to wait until the 16-s interval expired time before being presented with the memory questions. To keep the time given for the task completion consistent, we put a restriction of 11.3 s on the single-task processing and both single- and dual-memory tasks. However, unlike the processing tasks, participants could move forward as soon as they finished each memory task (for an illustration, see Figure 1).

Each trial started with a fixation point for 1,000 milliseconds. After the disappearance of the fixation point, there was a 1,000-ms delay. In the single-task processing block, participants received an expression evaluation to solve within 11.3 s. In the single-task verbal memory blocks, we presented a sequence of six words simultaneously, for a total of 5 s, followed by a 16-s blank screen, and then the memory list appeared; participants had 11.3 s to determine whether the list order was the same or different as the original order for that trial. Likewise, in the single-task spatial memory block, an arrangement of six crosses (i.e., reference memory array) was on the screen for 2 s, followed by a blank screen for 16 s, and then another set of six crosses was presented. Participants had 11.3 s to determine if the current solution matched with the actual answer for that trial. They could choose "yes" or "no" as an answer and move to the next trial after their response.

Each trial in both single-task memory blocks ended with feedback (Correct, Incorrect) for 1,000 milliseconds after a pause of 200 milliseconds. In dual-task blocks, after seeing the reference memory array (i.e., spatial arrangement or word chain), participants received the expression evaluation question to solve within 11.3 seconds. Then, 16 s after the disappearance of the reference arrays, they performed the memory tasks within 11.3 s by answering if the order of the verbal sequence (six words presented vertically) or spatial arrangement (six crosses on the square frame) matched the one shown at the beginning of the trial.

Data Analytic Approach

Accuracy was the key outcome variable. We used Cohen's d as an effect size measure. In the first of three analyses, we explored the effect of memory load on expression evaluation accuracy as a function of block type (i.e., expression evaluation, expression evaluation + verbal load, expression evaluation + spatial load). For the second and third analyses, we explored whether the expression evaluation task influenced verbal or spatial memory performance, respectively. We added the block type (i.e., for verbal memory performance, verbal load, verbal load + expression evaluation; for spatial memory performance, spatial load, spatial load + expression evaluation) as independent variables to examine if the processing task had a significant effect on memory recall. Finally, we conducted additional, follow-up analyses to learn more about the pattern of results, which will be explained in the Results sections of the experiments in which they are used.

For all analyses, we performed Bayesian instead of Frequentist inferential statistics. We did this because it is not possible to argue for the null hypothesis under frequentist testing (Jeffreys, 1961), but it possible with Bayesian analyses (Rouder et al., 2012), which has important implications for interpreting any null effects of memory loads on processing performance and vice versa (cf. Guitard et al., 2020). We determined Bayes Factors (BF) for binary logistic regression through "brms" package in R (Bürkner, 2017; R Core Team, 2019) to examine performance differences between load conditions using trial-level performance data. This method accounts for the binary distribution of our data (correct or incorrect) and includes guessing rate in the model (50% correct with two-forced choice). Such multilevel models allow modeling of data that take complex dependency structures into account and yield not only the mean but also a measure of the uncertainty of each parameter (the Bayesian Credible Interval). This model estimates the effect on the parameter η ('*eta*'; ability) by load-condition, using a Bernoulli distribution.

For all models, item and participant identity were included as random intercepts, to account for individual variation. The dependent variable was Correct versus Incorrect responses (1 or 0) on each trial ('Accuracy' in the example below). We used a normally distributed prior for eta (with $M = 0$, $SD = 5$, specified using: 'prior(normal(0,5)' below). The basic model including these parameters, as specified in *R*, is presented here:

$$
\begin{aligned}
&my\_model < -\, brm((bf(Accuracy \sim 0.50 + 0.50 \times \\
&inv\_logit(eta),\ eta \sim (1|ID) + (1|Item),\ nl = TRUE),\ data \\
&\quad = my\_data,\ family = bernoulli(\text{"}identity\text{"}),\ save\_all\_pars \\
&\quad = \ TRUE,\ prior \\
&\quad = prior(normal(0,\ 5),\ nlpar = \text{"}eta\text{"}),\ sample\_prior \\
&\quad = TRUE,\ iter = 10000,\ seed = 123)
\end{aligned}
$$

See Bürkner (2019) for details on *R* syntax and model specifications for Bayesian item response modeling.

We used a normally distributed prior for *eta*. For each model parameter, we report the parameter estimate, and its' 95% credible interval (CI). To obtain the BF in favor of the model including the load condition, we examined posterior distributions for the parameters, as well as compared the models including the load condition and individual participant effects, with a model with only the individual participant effects. We used the '*bayes_factor()*' function which approximates the model's marginal likelihood using the bridge-sampling algorithm, see Gronau et al. (2017, 2020) for further details.

BF provides information about the probability of the data given the presence of an effect (i.e., alternative hypothesis) compared with the probability of data given the absence of an effect (i.e., null hypothesis). For instance, a BF of 10 means that the likelihood of the alternative hypothesis is 10 times the likelihood of the null hypothesis based on the present data (Wetzels & Wagenmakers, 2012). There are different views on how to choose thresholds and name the intervals between them (e.g., Jeffreys, 1961; Kass & Raftery, 1995; Schönbrodt & Stefan, 2018; Schönbrodt et al., 2017; Schnuerch et al., 2021). In the current study, the criteria for the magnitude of the effects were based on a combination of terms for thresholds proposed by others, including Jeffreys (1961), Lee and Wagenmakers (2014), Schönbrodt and Stefan (2018), and van Doorn et al. (2021). We used category labels in which a Bayes Factor for the alternative hypothesis below 3 is inconclusive, 3–10 = moderate evidence, 10–30 = strong evidence, 30–100 = very strong evidence, and 100+ = decisive evidence for the alternative hypothesis. Conversely, the reciprocal of each number expresses support for the null, so, for example, a BF below 1/10, or .10, indicates positive support for the null hypothesis. When the outcome of this model comparison is 'inconclusive' (defined as a BF between .33 – 3), we use a directional test to compare the hypothesis that the load effect was larger than 0, as opposed to it being smaller than 0.

Our models seemed to converge well, as indicated by a r^ value of 1, and visual inspection of the parameter trace plots, showing random scatter around a mean value.

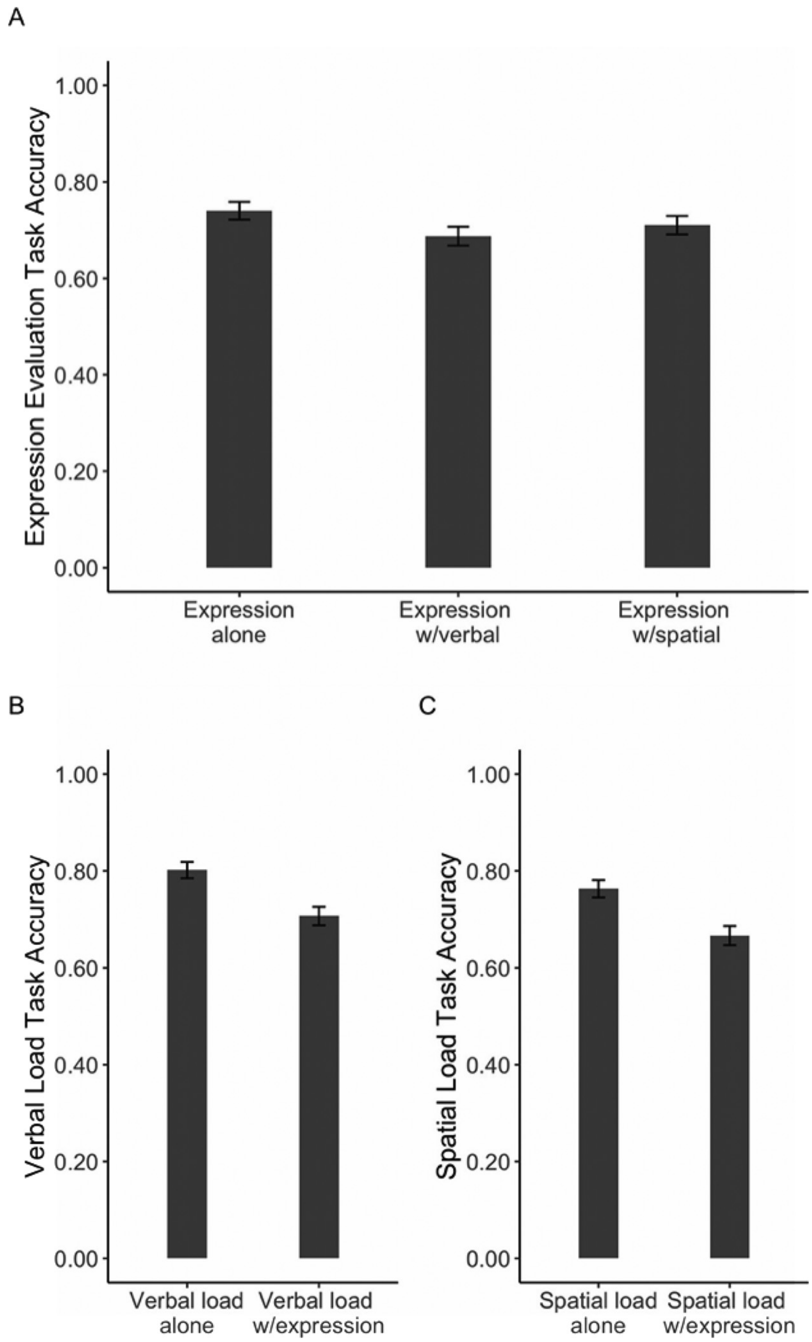
In all models below, the single-load (baseline) condition was used as the cornerstone condition, to which performance in the load condition(s) was compared. A negative value of η indicates that performance was poorer in the specified load condition, whereas values close to 0 indicates no difference. For each model, the CIs (the values in square brackets) indicate the lower and upper bounds of the 95% CI of the posterior distribution for the parameter, indicating that given the data and our prior assumptions, there is a .95 probability that this interval encompasses the effect of η.

Results

As shown in Figure 2, the expression evaluation accuracies seem to differ slightly across conditions, and the differences disappear when individual differences are considered. The statistical analyses moderately support our visual inspection with a note that BF values were very close to .10, a threshold for the strong evidence for the null hypothesis. Processing task (i.e., expression evaluation) performance (baseline, $M = .74$, $SD = .20$) did not change in the presence of verbal memory load ($M = .69$, $SD = .20$; $d = .12$, $\eta = -.41$, $SE = .34$, 95% CI [−1.07, .26], BF = .15). Similarly, there was no difference between performance in the baseline processing task and the spatial memory load condition ($M = .71$, $SD = .19$; $d = .07$, $\eta = -.36$; $SE = .31$, 95% CI [−.98, .25], BF = .12). A separate contrast suggested that expression evaluation performance did not differ between verbal and spatial memory load, as well (BF =.07).

In contrast, evaluating expressions disrupted both verbal memory (Verbal, $M = .80$, $SD = .19$; Verbal with expressions, $M = .71$, $SD = .19$) and spatial memory (Spatial, $M = .76$, $SD = .17$; Spatial with expressions, $M = .67$, $SD = .17$). The results were strong for both verbal memory ($d = .22$, $\eta = -1.16$, $SE = .31$, 95% CI [−1.79, −.59], BF = 228.5) and spatial memory ($d = .21$, $\eta = -.99$; $SE = .30$, 95% CI [−1.60, −.43], BF= 19.9). There is no evidence for an interaction, indicating that the load effect was similar for the verbal and spatial memory tasks (BF = .08).

Figure 2
Graphs of Experiment 1



Note. (A) The effect of memory loads on expression evaluation accuracy. (B) The effect of expression evaluation on verbal memory recall. (C) The effect of expression evaluation on spatial memory recall. Error bars represent the standard errors of the means.

Discussion

The results of Experiment 1 revealed that maintaining verbal or spatial items in memory did not appear to affect the accuracy of evaluating expressions. Conversely, the processing algebra expressions compromised both verbal and spatial memory loads. The latter supports possible spatial and verbal processing in expression evaluation. However, due to well-documented association between working memory and arithmetic (for a review, see DeStefano & LeFevre, 2004; Peng et al., 2016), the roles of spatial and verbal working memory in expression evaluation above and beyond arithmetic are unclear. Thus, next, we carried out Experiment 2 using arithmetic instead of algebraic expressions, for a comparison with Experiment 1.

Experiment 2: Arithmetic Processing

Method

Participants

We recruited 60 participants through Prolific and dropped two of them for their existing or potential distraction statements, leaving a total of 58 participants (21 Females, 37 Males). The mean age of the participants was 24.08 years (SD = 3.66, range = 18.4–30.9). All except one (2%) of the participants were native English speakers, and majority of them (86%) had at least a college degree.

Seven percent of participants were Hispanic or Latino, and 2% preferred not to report their ethnicity. Their racial distribution was 62% White or European, 19% Asian, 10% Black or African American, 2% American Indian/Alaskan, 2% other, 3% multiracial, and 2% unreported. Regarding devices, 2% and 98% of participants used small and large screens, respectively. All of them had normal or corrected vision.

Materials

For memory tasks, we presented the same six-word chains and spatial arrangements used in Experiment 1. For the arithmetic processing task, we used the same equations as in the previous experiment and replaced the variables with the associated numerical values. Therefore, the primary difference across Experiment 1 and the current experiment was the absence of unknown values. We presented the arithmetic equations in the form of expression evaluation items and added parentheses to reduce order of operations confusions [for example, expression evaluation form, $6y - x = 64$, given $x = 9$, $y = 13$; arithmetic form, $(6 \times 13) - 9 = 64$].

Procedure

The procedure was the same as in the previous experiment. The only difference was the absence of the presentation of the x and y values, which took 2.5 s in the previous experiment. To keep the duration between the memory array and memory question constant across experiments (i.e., 16 s), we increased the interval between the initial and final presentation of the memory items (e.g., word list) from 11.3 s to 13.8 s.

Analysis

We followed the same procedures as previous experiments.

Results

As shown in Figure 3, participants' mean scores for arithmetic equations were similar across all conditions (Arithmetic, M = .82, SD = .14; Arithmetic with verbal, M = .79, SD = .18; Arithmetic with spatial M = .82, SD = .19). Arithmetic performance did not differ under verbal (d = .08, $\eta = -.24$, SE = .26, 95% CI [−.75, .27], BF = .08) or spatial memory load (d = .01, η = .05, SE = .26, 95% CI [−.45, .55], BF = .05), compared with the baseline condition (i.e., arithmetic alone). A separate contrast confirmed with moderate evidence that arithmetic performance did not differ between verbal and spatial memory load (BF = .17).

Also, we found moderate evidence that solving the arithmetic problems had little effect on verbal memory (Verbal, M = .81, SD = .18; Verbal with arithmetic, M = .77, SD = .19; d = .11, $\eta = -.47$, SE = .41, 95% CI [−1.27, .33], BF = .17). However, solving the arithmetic problems substantively disrupted spatial memory (Spatial, M = .76, SD = .19; Spatial with arithmetic, M = .68, SD = .15; d = .20, $\eta = -2.34$, SE = 1.02, 95% CI [−4.69, −.75], BF= 23.8). The interaction of the modality with the presence or absence of load was, however, inconclusive (BF = .42), suggesting that the difference between load effects in the verbal versus spatial modalities still must be viewed cautiously.

Discussion

The results of Experiment 2 showed that there was no effect of memory loads on the processing of arithmetic expressions. However, the presence of the processing task appears to have differentially affected spatial and verbal memory performance. Although it was detrimental to spatial recall, there was apparently no effect on verbal recall. The effect on spatial recall was like that found in Experiment 1, but the role of verbal working memory clearly differed between arithmetic and algebraic expression evaluation.

The evidence supports the existence of verbal processes that uniquely belong to the processing of algebraic variables but not arithmetic expressions. However, this result may not be generalizable to all aspects of algebra. In Experiment 3, we used a coordinate plane task in which simple algebraic assignments that also involved variables were used to define a point on the plane. This is a form of algebra that does not include arithmetic calculation. The key question is whether verbal working memory processes will still play a role under these conditions, and at the same time assess the expectation that spatial working memory will play a role.

Experiment 3: Coordinate Plane

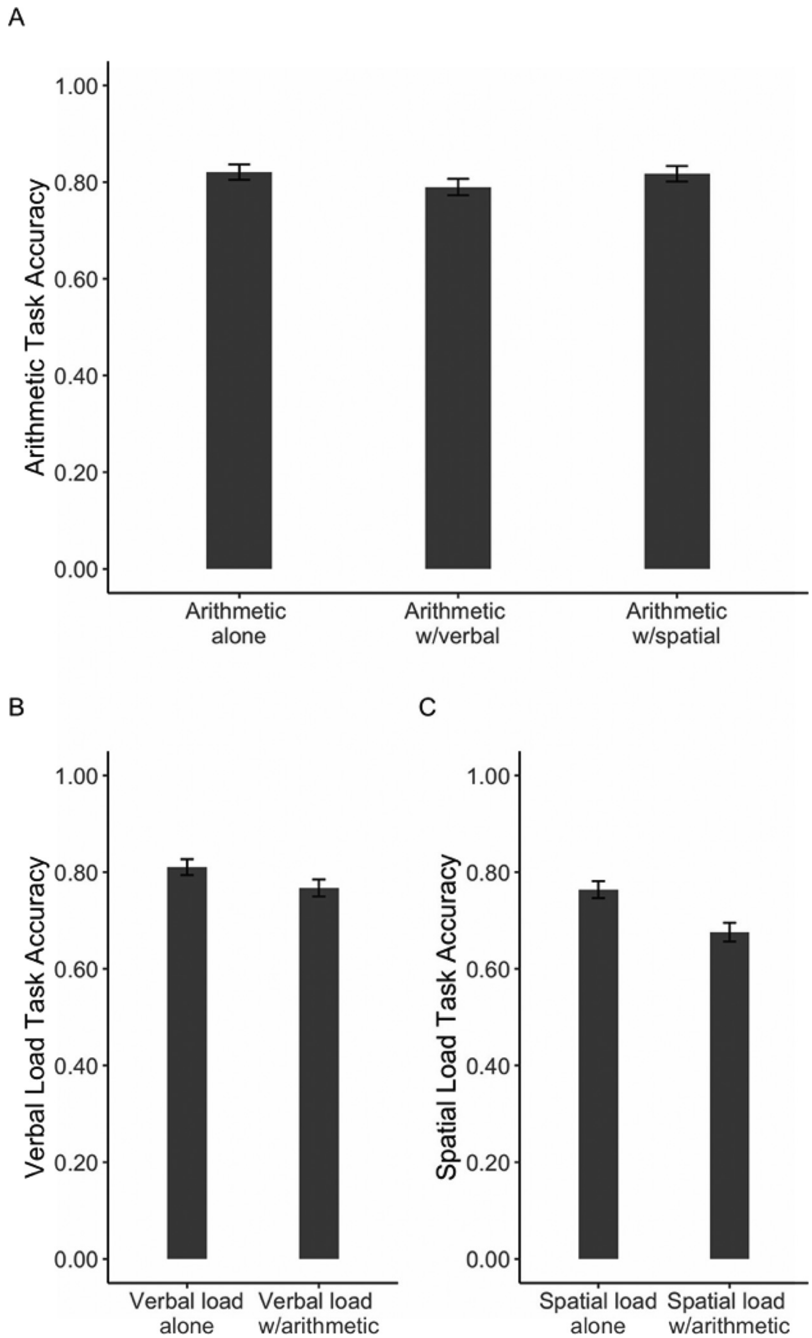
Method

Participants

We initially recruited 60 participants through Prolific but excluded seven of them due to their existing or potential distraction statements, resulting in 53 participants (31 females and 22 males) with a mean age of 24.18 years (SD= 4.35, range = 18.1–39.4). All were native English speakers and most of the participants (85%) held at least one college degree.

Ten percent of the participants were Hispanic, or Latino/Latina. Their racial distribution was 72% White or European, 19% Asian, 5% Black or African American, and 2% multiracial; 2% did not report their race. Regarding devices, 25% and 2% of participants used a small screen and a tablet, respectively, and 73% used a large screen. All had either normal or corrected vision.

Figure 3
Graphs of Experiment 2



Note. (A) The effect of memory loads on arithmetic problem-solving accuracy. (B) The effect of arithmetic problems on verbal memory recall. (C) The effect of arithmetic problems on spatial memory recall. Error bars represent the standard errors of the means.

Materials

For memory tasks, we used the same six-word chains and spatial arrangements. In the coordinate plane processing task, participants determined (yes, no) whether the location of a given point was in the correct position (or not) for the presented ordered pairs (i.e., algebraic variable values, e.g., $x = 30$, $y = -40$). The ordered pairs consisted of numbers that were multiples of ten (range 10–90). As previously mentioned, each block contained ten trials. We prepared three sets of ten items for each block (single task, verbal load, spatial load). For half of the 10 questions, the point on the coordinate plane matched the previously shown ordered pair, and for the other half it did not (i.e., a 20 unit up-down or left-right change, e.g., ordered pair: 20, −30 and point location: 20, −50; or flipped on the x- or *y*-axis, e.g., ordered pair: 20, −30 and point location: -20, −30).

Procedure

The procedures were identical to those described for in Experiment 1 (i.e., expression evaluation), except that the processing task involved coordinate plane (see Figure 4).

Analysis

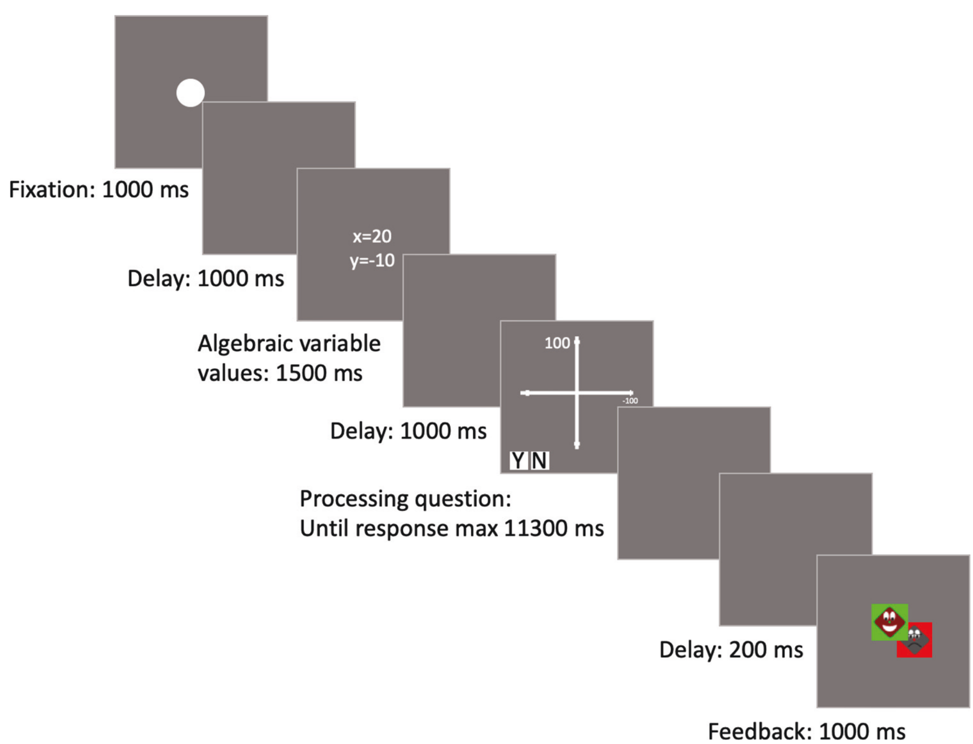
We followed the same procedures as in the previous experiments.

Results

As shown in Figure 5, the presence of the memory load had a noticeable effect on coordinate plane performance. Coordinate task performance decreased under spatial memory load (Coordinate, $M = .85$, $SD = .17$; Coordinate w/spatial, $M = .78$, $SD = .18$; $d = .18$, $\eta = -.93$, $SE = .27$, 95% CI [−1.57, −.33], BF = 6.0). The difference between coordinate plane performance with and without verbal load was within the indeterminate range (Coordinate w/verbal, $M = .80$, $SD = .18$; $d = .12$, $\eta = -.67$, $SE = .29$, 95% CI [−1.25, −.12], BF= .97). However, a directional test indicated that given the model and the data, it was 117.3 times more likely that verbal load would impair coordinate plane performance relative to improve performance. Bayesian logistic regression analysis was consistent with our visual inspection also. Coordinate plane performance had similar performance levels under verbal and spatial memory load (BF =.06), and both were numerically smaller than the coordinate task with no load. In sum, there is evidence for a load effect, though it is still tentative for a verbal load.

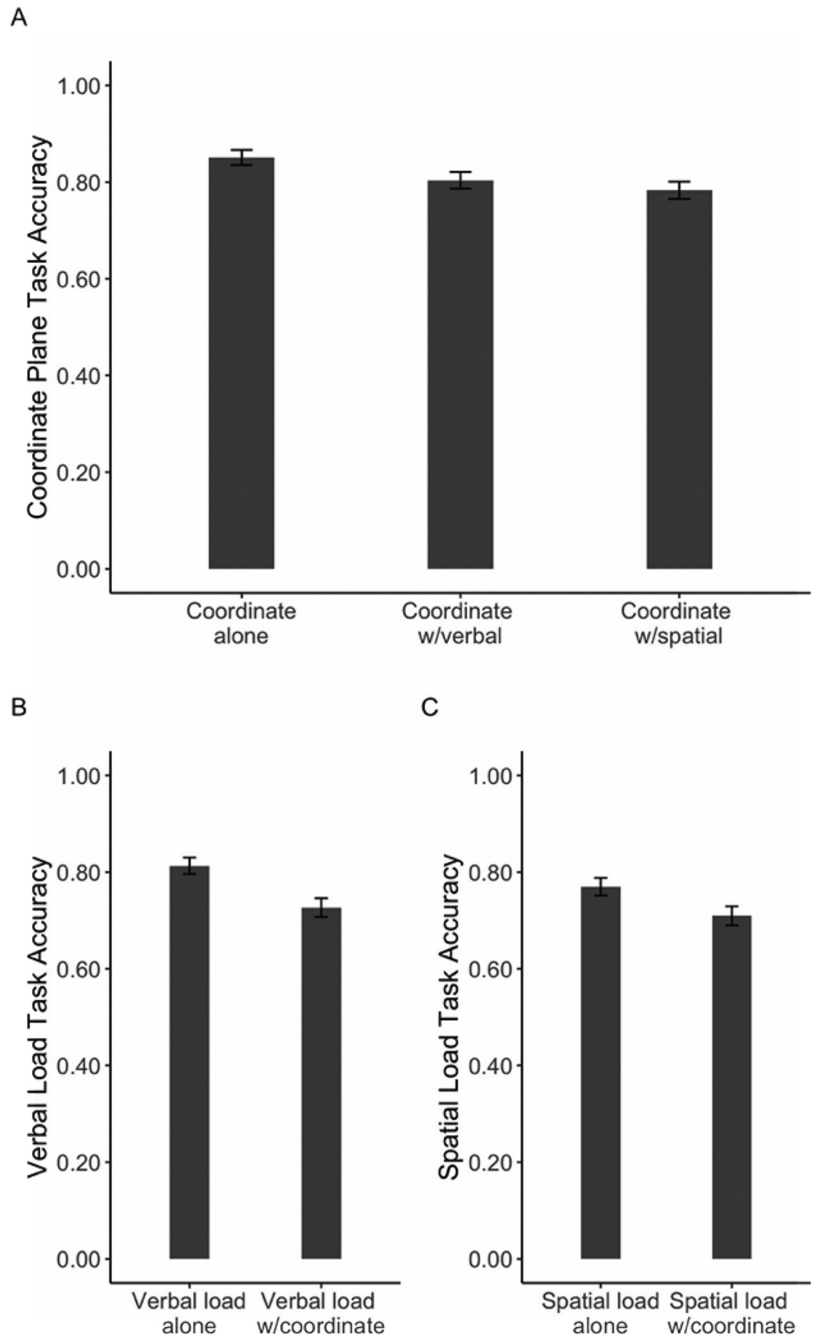
Engaging in the coordinate plane task markedly lowered verbal memory accuracy (Verbal, $M = .81$, $SD = .18$; Verbal with coordinate, $M = .73$, $SD = .18$; $d = .21$, $\eta = -1.03$, $SE = .29$, 95% CI [−1.62, −.48], BF = 65.00), but the effect on spatial memory (Spatial, $M = .77$, $SD = .18$; Spatial with coordinate, $M = .71$, $SD = .18$) was inconclusive ($d = .14$, $\eta = -.76$, $SE = .31$, 95% CI [−1.39, −.18], BF = 1.6). Given the model and the data, it was 209.5 times more likely the effect was smaller than 0, indicating that the load impaired memory performance. There was no evidence for an interaction of the modality with the presence or absence of load (BF = .09), suggesting that the load effect was similar for the verbal and spatial memory tasks. Overall, there is strong evidence for a load effect but in this case, in contrast to the findings for the processing task, the evidence is stronger for verbal than for spatial memory.

Figure 4
Example of Coordinate Plane Task in Experiment 3



Note. See the online article for the color version of this figure.

Figure 5
Graphs of Experiment 3



Note. (A) The effect of memory loads on coordinate plane accuracy. (B) The effect of coordinate plane on verbal memory recall. (C) The effect of coordinate plane on spatial memory recall. Error bars represent the standard errors of the means.

Discussion

The Experiment 3 results revealed that holding spatial information in WM reduced accuracy on the coordinate plane task. At the same time, the coordinate plane task impaired verbal memory recall substantially. The findings reveal the role of verbal working memory in coordinate plane processing, suggesting a unique role of verbal WM in processing algebraic variables. However, it was not certain whether the effect was solely attributable to a need to retain the *x* and *y* values during the coordinate plane task. To clarify this issue, Experiment 4 included a geometry task that does not include variables and thus should not require verbal processing.

Experiment 4: Geometry

Method

Participants

Initially, 60 participants completed the experiment through Prolific. We removed one due to an existing or potential distraction statement or lower performance, which yielded 59 participants (26 females, 15 males, and two other). Their mean age was 24.14 years (*SD* = 3.82, range = 18.3–30.8), and 85% reported completion of at least one college degree. All participants were native English speakers.

Two percent were Hispanic or Latino/Latina, two percent did not report their ethnic identity. Their racial distribution was 69% were White or European, 17% Asian, 7% Black or African American, 2% American Indian, and 3% multiracial; 2% did not report their race. Ten percent of the participants used a small screen, and 2% used a tablet and 88% a large screen. All had either corrected or normal vision.

Materials

The six-word chains and spatial arrangements were the same as in the previous experiments. For geometry processing task, we used parallel and nonparallel lines (for an example, see Appendix B). Each item included at least five lines, and at least two of them were parallel. We marked two angles between the lines, and participants determined whether the two angles were the same or not. The correct answers were divided 50/50 between “the same” and “different.”

Procedure

The procedures were identical to those described for Experiment 2 (i.e., arithmetic), except that the arithmetic task was replaced by a geometry task.

Analysis

We used the same procedure as the previous experiments.

Results

Surprisingly, we found moderate evidence that geometry performance increased under verbal memory load (Geometry, *M* = .79, *SD* = .19; Geometry with verbal, *M* = .85, *SD* = .16; *d*=.15, η = .78, *SE* = .28, 95% CI [.24, 1.36], BF = 3.25), compared with single-task geometry performance. The BF result showed that the difference between single-task geometry performance and geometry performance under spatial memory load was not large (*M* = .83, *SD* = .16, *d* = .11, η = .43, *SE* = .25, 95% CI [−.07, .92], BF = .22). As shown in Figure 6, participants’ geometry performances under verbal and spatial load were similar (BF = .09), although the evidence was stronger, though still moderate, in the case of a verbal load.

Engaging in the geometry task had no effect for verbal (Verbal, *M* = .81, *SD* = .21; Verbal w/geometry, *M* = .78, *SD* = .20; *d* = .08, η = −.56, *SE* = .41, 95% CI [−1.39, .26], BF = .22) or spatial (Spatial, *M* = .73, *SD* = .16; Spatial with geometry *M* = .68, *SD* = .15, *d* = .11, η = .25, *SE* = .50, 95% CI [−.81, 1.18], BF = .11) memory accuracy. The BF result moderately favored the null hypothesis. There was moderate evidence against an interaction, meaning that the load effect was similar for the verbal and spatial memory tasks (BF = .18).

Discussion

The Experiment 4 results suggested that there is an effect of at least a verbal memory load on the geometry task but that, surprisingly, it is a positive effect. If this effect is replicated, it might occur because the geometry task makes little demand on working memory and the presence of a memory task helps to keep participants engaged or alert (for a similar finding developmentally see Cowan et al., 2021). This account also is compatible with the finding of an effect of geometric processing on memory.

To clarify the role of task difficulty, in Experiment 5 we examined the process of mental rotation, which is quite demanding but does not rely on verbal processing (Smyth & Scholey, 1994) and thus no effect of processing on verbal recall, but strong detrimental effect on spatial recall is expected.

Experiment 5: Mental Rotation

Method

Participants

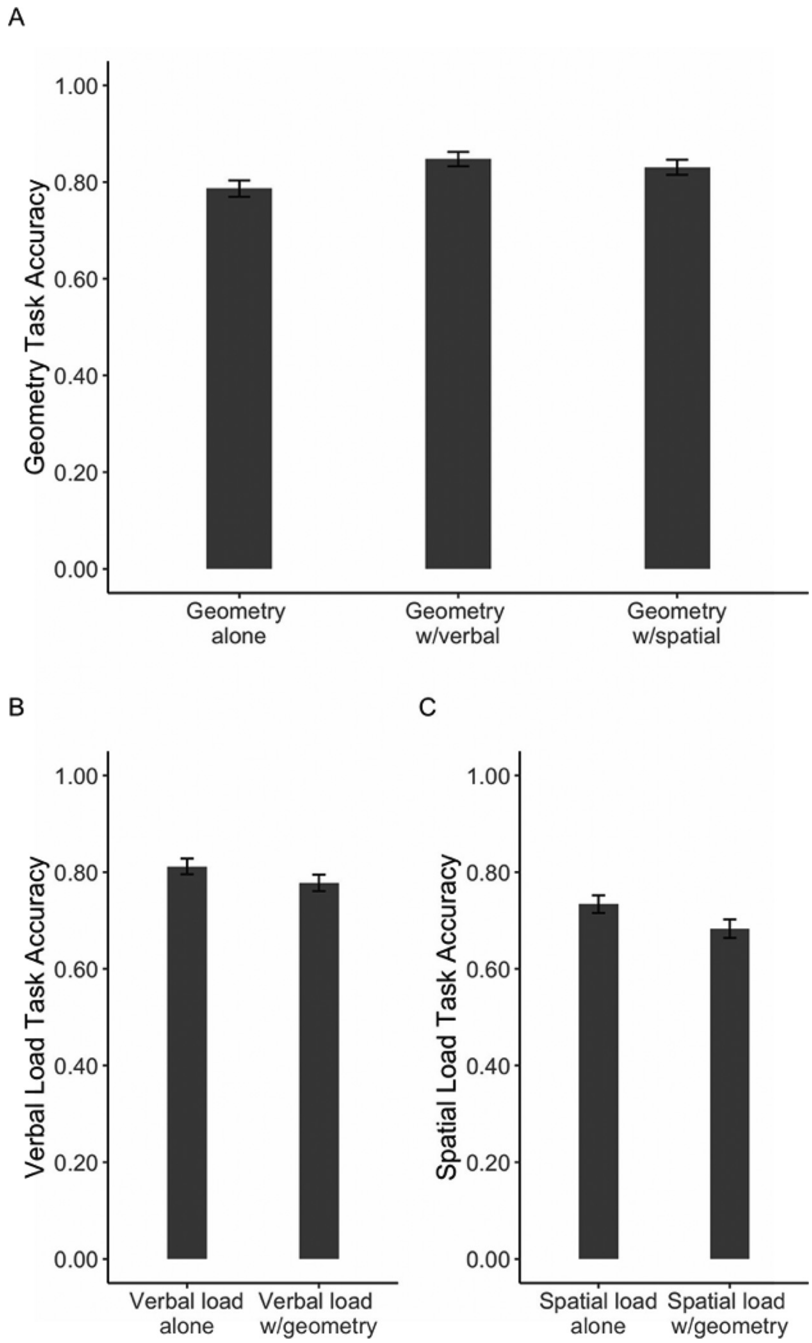
Sixty participants completed the experiments through Prolific. We removed the data of eight participants due to an existing or potential distraction statement, resulting in 52 participants (23 Females, 29 Males). The mean age of these 52 participants was 22.68 years (*SD* = 3.26, range 18.11–30.00). All except one (98%) were native English speakers, and most of whom (81%) had at least one college degree.

Four percent of participants were Hispanic or Latino, and 4% of participants preferred not to declare their ethnicity. Their racial distribution was 77% White or European, 15% Asian, 2% Black or African American and 4% multiracial; 2% did not report their race. Six percent used a small screen, 2% a tablet, and 92% a large screen. All participants had a normal or corrected vision.

Materials

We prepared the drawings for the mental rotation questions using the National Council of Teachers of Mathematics’ (NCTM) Isometric Drawing Tool. There were two drawings on the screen, the target figure and its pair (see Appendix C). Each target figure was a three-dimensional shape containing ten small cubes. We constructed the paired figure using one of the following rotation angles depending on whether it matched the target figure: (1) 75, 60, or 90°, along both *x* and *y* axis; 120 or 150 along *x* axis only (if it matched the target figure), and (2) 30° along *x* and *y* axis; 75, 90, 120, or 150° along the *x* axis (if it did not match the target figure). To create difference pairs, we changed the location of two blocks in the original shape and then rotated it.

Figure 6
Graphs of Experiment 4



Note. (A) The effect of memory loads on geometry accuracy. (B) The effect of geometry on verbal memory recall. (C) The effect of geometry on spatial memory recall. Error bars represent the standard errors of the means.

Procedure

The procedures were identical to those described for the previous experiment, except that the geometry task was replaced by a mental rotation task.

Analysis

We used the same procedure as in previous experiments. However, we added gender as a covariate due to well-replicated sex differences in the mental rotation task (for a review, see Voyer et al., 1995). The sex-related results are in the additional materials.

Results

As shown in Figure 7, mental rotation task performance (baseline) was somewhat lower under verbal memory load, however; the Bayes Factor in favor of the model was in the indeterminate range (Mental rotation, *M* = .72, *SD* = .13; Mental rotation with verbal load, *M* = .65, *SD* = .17; *d* = .15, η = −.75, *SE* = .32, 95% CI [−1.40, −.15], BF = 1.26). A directional test indicated that given the model and the data, it was 145.0 times more likely that verbal load would impair mental rotation performance instead of improving it. On the other hand, we found moderate evidence that no differences between mental rotation task performance and performance under spatial memory load (*M* = .75, *SD* = .15; *d* = .07, η = .34, *SE* = .27, 95% CI [−.17, .87], BF = .12). Yet, mental rotation performance differed between verbal and spatial memory load (BF =120.83). In sum, there may be slight effects of the two loads in opposite directions, with any detrimental effect of load occurring in the case of the verbal load.

Engaging in mental rotation processing had no effect on verbal memory recall, as shown in Panel A of Figure 7 (Verbal, *M* = .79, *SD* = .18; Verbal with mental rotation, *M* = .80, *SD* = .20); in fact, the comparison of the verbal memory recall with and without mental rotation conditions yielded strong evidence for the null, (*d* = .01, η = .11, *SE* = .25, 95% CI [−.38, .62], BF= .06). However, as shown in Panel C of Figure 7, engaging in the mental rotation task substantially disrupted spatial memory recall (Spatial, *M* = .77, *SD* = .18; Spatial with mental rotation, *M* = .68, *SD* = .14; *d* = .07, η = −1.61, *SE* = .40, 95% CI [−2.46, −.90], BF = 7307.2). There was strong evidence for an interaction, meaning that the load effect was different for the verbal and spatial memory tasks (BF = 91.82).

Discussion

results of Experiment 5 revealed that the verbal and spatial load did not substantively compromise mental rotation performance (with a possible but unproven detrimental effect of verbal load) and engaging in the mental rotation task apparently had no effect on verbal memory. In contrast, spatial load was severely compromised by mental rotation processing. The results, in combination with those of the previous experiments, appear to support an attentional component of spatial working memory, independent of the task nature (i.e., spatial and verbal), directly affected by the task difficulty. To assess the latter, in the next section we carried out additional analyses to compare results across experiments. First, we report cross-experimental comparisons regarding difficulty levels. Then we report cross-experimental results from a postexperimental questionnaire on self-reported strategies.

Cross-Experiment Comparisons of Difficulty Levels

To evaluate our hypothesis that processing-task difficulty influenced our results, we adopted a measure that enabled a comparison of difficulty levels across experiments. The measure follows the Inverse Efficiency Score proposed by Townsend and Ashby (1978). That score takes the mean reaction time (RT) for correct responses and divides it by the accuracy such that the RTs of less accurate participants were adjusted upward. We modified this adjustment to yield a theoretically more suitable measure that takes guessing into account: RT for trials in which the answer was known divided by the proportion of trials in which the answer was known, which we term *Difficulty Level*, as elaborated below.

Proportion of Knowing

Figure 8 shows a tree diagram indicating how we can estimate the proportion of trials in which the answer was known. The correct answers participants provided were assumed to comprise a mixture of responses based on knowing the correct response and answers based on guessing. Given that the correct answer could be guessed at a rate of .50, the correct responses can be decomposed into the proportions of known responses, p(know) and, among the remaining trials, the proportion of correct guesses, equal to [1-p (know)](.50):

$$\mathrm{p(correct)} = \mathrm{p(know)} + [1 - \mathrm{p(know)}] \times (.5) \tag{1}$$

This equation can be rearranged to yield an estimate of the desired quantity, p(know):

$$\mathrm{p(know)} = [\mathrm{p(correct)} - .50]/(.50) \tag{2}$$

RT on Know Trials

RTs for correct responses also can be decomposed into trials in which participants knew the answers and the others in which they guessed correctly. Where *N*(know) is the number of correct trials when the answer was known and *N*(correct by guess) includes all remaining correct trials, *N*(know) and *N*(correct by guess) sum to *N*(correct). On these trials, respectively, the average RTs are RT(know) and RT(correct by guess). It can be assumed that RT(correct by guess) is about the same as RT for incorrect responses, and that *N*(correct by guess) is about the same as the number of incorrect responses, given that the correct-guessing rate is .50. Given that *N*(incorrect) is available from the data, it is substituted for *N*(correct by guessing). To estimate the average of RT on know trials, we need to consider the weighting of knowledge and guessing within RT(correct) according to the number of trials:

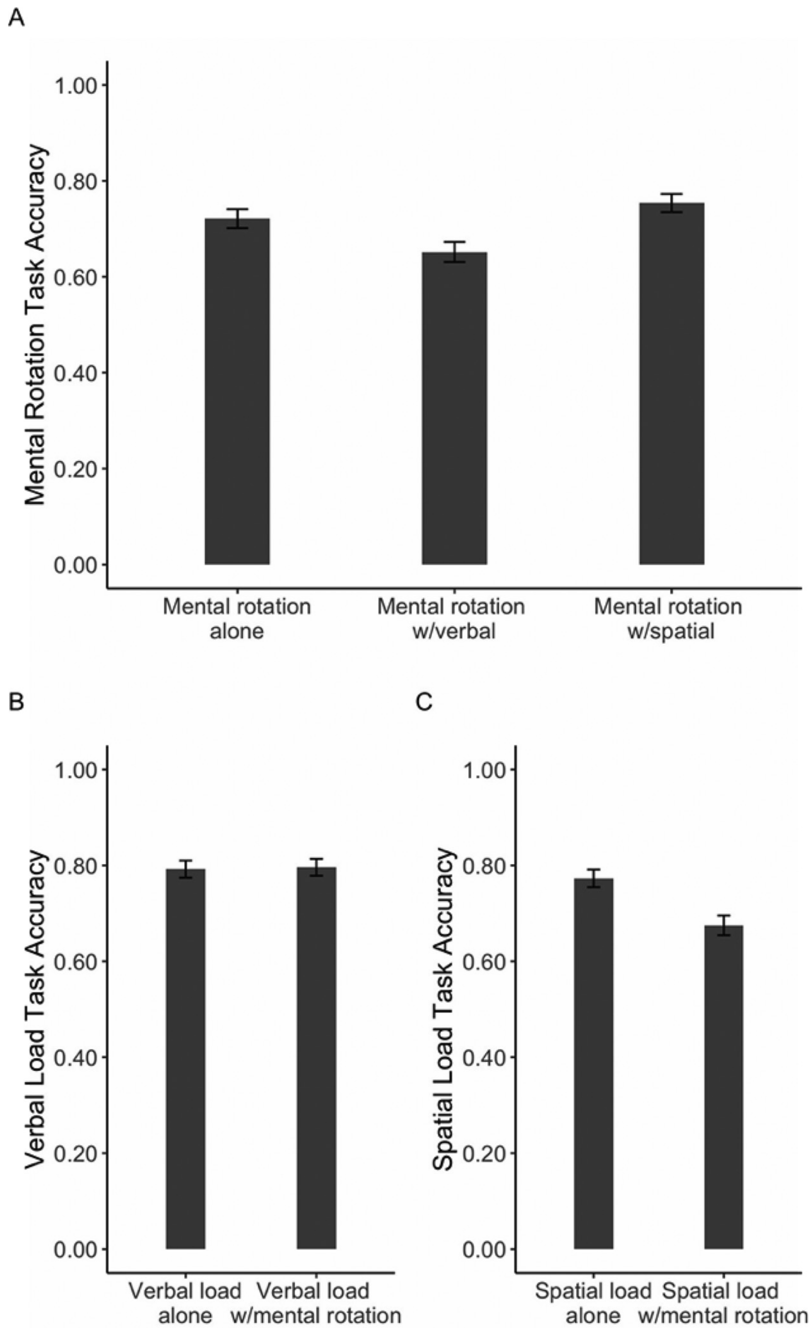
$$\begin{aligned}\mathrm{RT(correct)} = [&\mathrm{N(know)} \times \mathrm{RT(know)} \\ &+ \mathrm{N(incorrect)} \times \mathrm{RT(incorrect)}]/\mathrm{N(correct)}\end{aligned} \tag{3}$$

This formula can be rearranged to yield an intermediate result:

$$\begin{aligned}\mathrm{N(know)} \times \mathrm{RT(know)} = [&\mathrm{RT(correct)} \\ &\times \mathrm{N(correct)} - \mathrm{N(incorrect)} \times \mathrm{RT(incorrect)}]\end{aligned} \tag{4}$$

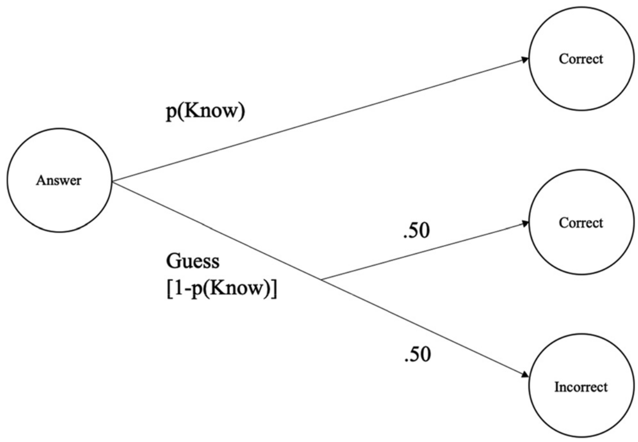
We also know that *N*(know) = *N*(correct) − *N*(correct by guess) = *N*(correct) − *N*(incorrect), which allows us to derive the desired quantity, RT(know):

Figure 7
Graphs of Experiment 5



Note. (A) The effect of memory loads on mental rotation accuracy. (B) The effect of mental rotation on verbal memory recall. (C) The effect of mental rotation on spatial memory recall. Error bars represent the standard errors of the means.

Figure 8
Demonstration of Kinds of Answers



$$\text{RT}(\text{know}) = [\text{RT}(\text{correct}) \times \text{N}(\text{correct}) - \text{N}(\text{incorrect}) \times \text{RT}(\text{incorrect})]/[\text{N}(\text{correct}) - \text{N}(\text{incorrect})] \quad (5)$$

Alternatively, the numerator and denominator on the right-hand side of the equation can both be divided by the total number of trials in the experiment to yield the following:

$$\text{RT}(\text{know}) = [\text{RT}(\text{correct}) \times \text{p}(\text{correct}) - \text{RT}(\text{incorrect}) \times \text{p}(\text{incorrect})]/[\text{p}(\text{correct}) - \text{p}(\text{incorrect})] \quad (6)$$

Overall Difficulty Formula

We divided average RT(know) by probability of knowing, p (know), as an index of the difficulty level of each task.

$$\text{Difficulty level} = \text{RT}(\text{know})/\text{p}(\text{know}) \quad (7)$$

As shown in Table 2, comparisons of task difficulties in the different experiments yield the following order: MRT > Expression Evaluation > Arithmetic > Geometry > Coordinate Plane. We also calculated the mean task difficulty across all processing tasks for further demonstration and then found each task's relative position to this average. As shown in Table 3, geometry and coordinate plane were easier than the other tasks. Altogether, these results confirm our speculations about the relative difficulty levels in the different experiments.

Cross-Experiment Comparisons of Memory Maintenance Strategies

Last, we examined whether the strategies of holding memory items change across different experiments, which could be crucial for accurately interpreting the effects of various processing tasks on the same memory loads. To do that, we asked participants to choose which strategies they used to hold verbal items in their minds while performing the processing task (e.g., expression evaluation) at the end of the experiment. The choices were:

1. I tried to use part of each word.
2. I tried to make a study out of the words.
3. I tried to group the words.
4. I tried really hard to remember.
5. I said the words to myself.
6. I just tried hard to remember.
7. I guessed.

Similarly, for spatial item recall strategies, we presented several choices. That includes:

1. I tried to separate the items into smaller clusters or groups.
2. I tried to see what objects the crosses looked like.
3. I tried to make patterns.
4. I guessed.
5. I gave up.

Each item has five options, namely—Never (1 pts), Rarely (2 pts), Sometimes (3 pts), Often (4 pts), and Always (5 pts). For each item under each condition (verbal and spatial), we averaged participants' scores. For instance, as shown in Table 4, the mean of "I tried to group words" is 3.0 for the mental rotation experiment, meaning that participants reported that they *sometimes* used this strategy while holding verbal items in their minds in that particular experiment. The mean score of each choice for both kinds of memory items was similar across all experiments (for graphs, see additional materials).

Summary of Cross-Study Comparisons

Our goal was to investigate the role of domain-general attention, as well as domain-specific working memory processes in algebraic tasks, including those that are symbolic (i.e., for expression evaluation) and spatial (i.e., for the coordinate plane). We also included associated nonalgebraic processing tasks to more accurately interpret our algebra-related experimental findings.

Among all experiments, coordinate plane (BF = 6.0) and geometry (BF = 3.25) were the only processing tasks that appeared to be affected by memory recall despite being relatively less difficult tasks. Interestingly, coordinate plane but not geometry performance seemed to decrease under spatial load and geometry performance increased under verbal load. We suggest that for the former, the spatial and verbal features of the coordinate plane in combination (unique to coordinate plane among all processing tasks) might have interfered with the features of spatial memory, indicating feature interference between processing and memory tasks. In contrast,

Table 2
Difficulty Levels of the Processing Tasks

Task	Av. of RT(correct)	Av. of RT(incorrect)	Av. of RT(know)	Av. of p(correct)	Av of p(know)	Av. of difficulty
Expression E.	6,637.10	7,490.16	6,176.32	0.74	0.48	12,849
Arithmetic	6,626.06	7,081.65	6,498.69	0.82	0.64	10,132
Coordinate P.	3,712.46	4,527.77	3,539.31	0.85	0.70	5,043
Geometry	5,370.91	5,624.94	5,276.22	0.79	0.57	9,210
Mental Rot.	6,040.66	6,251.77	5,907.58	0.72	0.44	13,356

Note. E = evaluation; P = plane; Rot. = rotation; Av. = average; RT = reaction time; p = proportion.

verbal memory load could potentially facilitate geometry task performance because participants might have gotten bored when the task was presented alone because it was an easier task; participants could not do this under spatial load since spatial attention was too attention-demanding. Alternatively, if geometric thinking experienced interference when people tried to use verbal coding, the verbal load might have prevented that coding and allowed visual coding to predominate without interference (Brandimonte et al., 1992; Dodson et al., 1997).

Regarding the effect of processing tasks on memory load, the results, across experiments, suggest that the processing of the numerical values associated with variables, whether in the context of expression evaluation (symbolic), BF = 228.5, or the coordinate plane (spatial), BF = 65.0, engages verbal working memory. Moreover, as shown in Table 1, we observed a detrimental effect on visuospatial load in the presence of the processing tasks selectively. Even though the arithmetic and expression evaluation tasks did not have any spatial features, we observed detrimental influences of these tasks on spatial memory (arithmetic, BF = 23.8; expression evaluation, BF = 19.9). However, spatial recall performance was not substantially affected by coordinate plane (BF = 1.6) or geometry (BF = .11) tasks, even though they included spatial components. The effects were apparently independent from the nature of processing tasks (i.e., spatial vs. verbal), however; the task difficulty levels (i.e., average of difficulties of expression evaluation, arithmetic, coordinate plane, geometry, and mental rotation; for details, see Table 2) contributed to differences in the influence of the processing tasks on spatial memory.

More challenging tasks (i.e., expression evaluation, arithmetic, and mental rotation) significantly impaired the maintenance of spatial load, whereas easier tasks did not substantively affect maintenance of the visual items. In other words, when participants completed more demanding processing tasks and then recalled the visual array, their retention of spatial information was compromised. The attentional demands of the more difficult tasks were too great to allow attention-sharing with spatial memory, which also depends on attention (e.g., Morey et al., 2013; Souza & Oberauer, 2017; Vergauwe et al., 2014). In this circumstance, attention switched from spatial memory maintenance to these processing tasks. Given this attention switch, there would be no interference on processing but a large effect of processing on spatial memory.

Table 3
Position to Mean Task Difficulty

Expression evaluation	0.818
Arithmetic	0.004
Coordinate plane	−1.521
Geometry	−0.272
MRT	0.971

Note. *M* task difficulty = 10,117.94, *SD* of task difficulty = 3,336.43. Formula = (Individual task Difficulty − *M* task Difficulty)/*SD*. The negative sign indicates the task is easier compared with the average task difficulty.

Last, our analysis of strategy uses for memory item maintenance across experiments revealed that participants reported similar techniques, regardless of the difficulty levels of the processing tasks; verbal rehearsal and pattern formation were the most commonly used methods to maintain word chains and spatial arrangements, respectively, in all experiments. These findings made our inferences even more robust, eliminating the possibility of the results attributable to changes in types of strategies used for item maintenance. We offer a theoretical account followed by a broader discussion of the role of working memory in algebra, implications for algebra, and implications for working memory.

General Discussion

As described earlier, Table 1 summarizes both the assumed task demands and the dual-task interference effects obtained in all experiments. We suggest that three concepts can organize the core findings. (a) Verbal or visual-spatial features in the memory load and processing tasks can conflict with one another when the features are similar, making it difficult to maintain the memory load while carrying out processing relying on similar features. (b) Visual memory loads draw more heavily on general attention than verbal loads, which can result in interference between a visual load and tasks that are difficult, even tasks that do not include an obvious visual-spatial processing component. (c) In dual-task procedures with strong conflict between storage and processing, it is usually the storage that suffers, with processing mostly protected. These results contribute to our understanding of memory and attentional systems broadly and make a unique contribution to our understanding of how these systems support algebraic processing and learning.

Effects of Processing on Storage of a Memory Load

Three processing tasks were assumed to impose verbal processing constraints: algebraic expressions (BF = 228.5), arithmetic (BF = .17), and coordinate plane placements (BF = 65.0). Of these, two produced a detrimental effect of processing on the verbal load, which did not seem to occur in the other experiment. The absence of a strong effect of processing on verbal load in the arithmetic task is likely because the problems were relatively easy,

Table 4
Average Strategy Scores for Memory Item Maintenance Across Five Experiments

Strategy	Exp *M* (*SD*)	Arth *M* (*SD*)	Coord *M* (*SD*)	Geo *M* (*SD*)	Mrt *M* (*SD*)
Verbal					
1. I tried to group the words.	3.1 (1.3)	3.0 (1.3)	2.9 (1.6)	2.8 (1.6)	3.0 (1.6)
2. I tried to make a story out of the words.	1.8 (1.1)	1.8 (1.1)	1.8 (1.1)	1.8 (1.3)	1.6 (1.2)
3. I said the words to myself.	4.2 (1.1)	4.0 (1.2)	4.5 (0.9)	4.1 (1.5)	4.4 (1.0)
4. I tried to use part of each word.	2.4 (1.5)	2.6 (1.7)	2.2 (1.5)	2.6 (1.6)	2.5 (1.6)
5. I just tried hard to remember.	3.4 (1.2)	3.0 (1.3)	3.6 (1.2)	3.7 (1.3)	3.4 (1.4)
6. I gave up.	1.2 (0.5)	1.2 (0.5)	1.2 (0.5)	1.2 (0.5)	1.2 (0.5)
7. I guessed	1.7 (0.7)	1.5 (0.6)	1.8 (0.7)	1.7 (0.8)	1.7 (0.7)
Spatial					
1. I tried to separate the items into smaller clusters or groups.	3.8 (1.2)	3.6 (1.3)	3.8 (1.1)	3.7 (1.3)	3.6 (1.3)
2. I tried to make patterns.	4.1 (1.0)	4.0 (1.3)	4.1 (1.1)	4.0 (1.1)	3.8 (1.1)
3. I tried to see what object the crosses looked like.	2.3 (1.4)	2.3 (1.4)	2.3 (1.2)	2.5 (1.6)	2.5 (1.5)
4. I tried really hard to remember.	3.2 (1.3)	3.2 (1.1)	3.8 (1.1)	3.8 (1.1)	3.6 (1.2)
5. I gave up.	1.2 (0.4)	1.2 (0.5)	1.2 (0.5)	1.4 (0.7)	1.3 (0.6)
6. I guessed	1.6 (0.7)	1.8 (0.9)	1.7 (0.8)	2.0 (0.9)	1.8 (0.8)

Note. *M* = mean; *SD* = standard deviation; Verbal = strategies used to maintain verbal memory load; Spatial = strategies to maintain spatial memory load; Exp = expression evaluation; Arth = arithmetic; Coord = coordinate plane; Geo = geometry; MRT = mental rotation.

compared with problems in some prior studies (Hitch, 1978), for college-educated adults. In this situation, much of the arithmetic processing was dependent on long-term (Geary et al., 1986) and not working memory (see, e.g., Doherty et al., 2019).

Three processing tasks were assumed to impose visual demands: coordinate plane placements, geometry, and mental rotation. However, the actual result was that a clear effect of processing on load among these tasks occurred only in the case of mental rotation (coordinate plane placements, BF = 1.6; geometry, BF = .22; mental rotation, BF = 7307.2). Moreover, evidence of an effect of processing on visual storage was obtained in two other experiments that were assumed to have no clear visual processing component (algebraic expressions, BF = 19.9, and arithmetic, BF = 23.8). Instead of a conflict with visual representations, interference with the visual load may occur because of attentional conflict between storage and processing, given the generally high attention demand of visual storage (e.g., Morey & Bieler, 2013) and our findings summarized in Table 2 that show a relatively high difficulty level for the algebraic expressions and arithmetic tasks.

One experiment showed only weak evidence of an effect on visual storage (by directional test), namely the coordinate plane task (BF = 1.6). However, this task was lowest in difficulty level and the dual-task conflict could have come not from attention, but from the visual demands of the task. The sole experiment showing no evidence of an effect on visual storage was the geometry task, which was rated relatively low in difficulty and presumably does not have a verbal component. Additionally, the mental rotation is the one that yields solid evidence for a distinct effect of the processing task on the verbal and spatial load performances, potentially due to an especially high attentional demand.

Effects of a Memory Load on Processing Accuracy

Across experiments, the effect of storage on the processing task was weak overall (Expression BF under verbal = .13, under spatial = .12; Arithmetic BF under verbal = .08, under spatial = .15; Coordinate plane under verbal = .97, under spatial = 6.00; Geometry under verbal = 3.25, under spatial = .22; Mental rotation under verbal = 1.26, under spatial = .12), in keeping with the previous results of Doherty et al. (2019) using memory for letters and an arithmetic processing task. This finding suggests that participants tended to drop the memory load instead of sharing attention between storage and processing (cf. Cowan et al., 2021).

Coordinate plane (BF = .97) and mental rotation (BF = 1.26) tasks showed weak evidence (by directional test) of lower performance in the presence of verbal processing tasks, suggesting that, in these tasks, verbal rehearsal or mnemonic processing could have persisted during the processing. Within these findings, an effect of verbal memory (but not spatial memory) on mental rotation would be our most anomalous result and might indicate that verbal memory is not completely attention-free. Indeed, under high-pressure situations, effects of verbal storage on visual processing can be observed, such as an effect of memory for digits on reaction times to press a button matching a signal location (Chen & Cowan, 2009). Future studies can elaborate on that by manipulating the difficulty level of mental rotation (e.g., using two-dimensional rotation) while keeping the verbal memory task the same, which would provide additional information on the interplay between verbal memory and mental rotation.

A detrimental effect of spatial load on processing seem to emerge only for the coordinate plane task (BF = 6.0), which produced no effects of processing on storage and was judged the least difficult task by far. Our suggestion is that, in this case, the spatial load conflicted with the visual requirements of the task, but participants still thought they could maintain the visual load during the task, resulting in visual feature conflict between the tasks. This does not appear to be an example of both tasks heavily engaging general attention.

Summary

The results, overall, suggest separate maintenance mechanisms for verbal and spatial items. Unlike spatial items-to-be-remembered, two discrete systems might exist to prevent verbal information decay; (a) attention-based and (b) verbal rehearsal/articulatory processing-based (i.e., use of language to maintain phonological information; Camos et al., 2011; Camos & Barrouillet, 2014). The idea of distinct maintenance mechanisms for spatial and verbal information is in line with the asymmetry hypothesis, which posits that the attentional demands of verbal and spatial item maintenance differ, with more attention needed in the spatial case. Verbal information can be sustained in working memory through articulatory processing, without very much attentional demand (Gray et al., 2017; Morey et al., 2013; Vergauwe et al., 2014).

Unlike most of the results, it could seem surprising to observe a cross-domain load effect on processing in the mental rotation experiment. When the verbal load exists, the accuracy of mental rotation (spatial in nature) appeared to drop. One possible explanation for that is the relatively higher attentional demand of the mental rotation task. It was attention-demanding enough that the active rehearsal mechanism for maintaining the verbal information interfered with the attentional mechanism, despite a limited, but nonzero, attentional requirement of rehearsal (Thalmann et al., 2019), causing a decline in the mental rotation performance. In contrast, participants would not be able to maintain and refresh the spatial load during mental rotation and may have only resumed mnemonic activities for the spatial load after completing the mental rotation task (cf. Cowan et al., 2021).

Algebra and Working Memory Processes

A relation between individual differences in working memory capacity and concurrent and longitudinal gains in mathematics achievement is well established (Adams & Hitch, 1997; Cragg et al., 2017; De Smedt et al., 2009; Geary et al., 2007). Moreover, experimental studies have revealed a relation between working memory processes and performance in arithmetic in adults (Hitch, 1978; Imbo & Vandierendonck, 2007), although this relation may vary with problem complexity and participants' problem-solving strategies (DeStefano & LeFevre, 2004). At the same time, little is known about the working memory demands of algebraic processing, despite the critical importance of algebra to students' mathematical development (NMAP, 2008).

Our results provide the first experimental evidence for both spatial and verbal processing for evaluating expressions and placing ordered pairs in the coordinate plane. Given that the latter was the least difficult task across the five experiments, a spatial WM effect for the coordinate plane could reflect storage capacity for visual inputs, as noted. However, an overall decrease in spatial load performance in the presence of expression evaluation which contains only verbal features (not spatial), might result from the attentional demands of visuospatial working memory. It is worth noting that individual differences in visuospatial memory growth across the elementary-school years predicts later mathematics achievement in middle school and high school better than developmental growth in verbal WM (Allen et al., 2020; Li & Geary, 2013; Li & Geary, 2017).

Our findings suggest that this pattern might come from the overlap, at least in part, in the domain-general attentional aspect of visuospatial working memory and the higher attentional demands of advanced mathematics. In other words, a domain-general attentional factor could be a core determinant of the link between visuospatial memory and mathematics, which is in line with current research implying that the focus of attention, rather than storage, is the driving factor in the relationship between working memory and higher-order cognitive skills (Gray et al., 2017).

We also obtained strong evidence for verbal processes in algebra that distinguishes it from arithmetic. These effects were found for both the expression evaluation and coordinate plane experiments. One common feature of these tasks that was not found in arithmetic processing was the need to encode and retain the numerical values of variables. These findings are in keeping with previous studies showing the processing demands of algebraic variables (Pollack et al., 2016; Pollack & Price, 2020), although we cannot determine whether their effects are attributable to interference from use of the same symbols (x, y) in reading or the demands of maintaining the numerical value of the variable in WM during processing (McNeil et al., 2010).

Either way, we also observed modality-related effects attributable to the distinct visual and spatial features of expression evaluation and the coordinate plane. Although spatial load did not seem to impede expression evaluation accuracy (BF = .12), as mentioned above, it appeared to compromise coordinate plane accuracy (BF = 6.00). To sum up, algebra appears to heavily need a domain-general attentional resource and modality-related verbal or visual resources depending on the algebra content. The former is mainly affected by task difficulty, whereas the latter relies more on the task features.

Implications for Algebra Learning

Learning, in general, occurs with the transformation of information from working memory to long-term memory. The transformation best occurs when the number of items represented in working memory does not exceed its limits, as shown in experimental studies (Cowan, 2014; Cowan et al., 2013) and in keeping with cognitive load theory (Paas et al., 2010; Van Merrienboer & Sweller, 2005). However, as demonstrated in our experiment results, algebra-related tasks are cognitively demanding, whether they primarily involve spatial (i.e., coordinate plane) or verbal (i.e., variables) processing, indicating that different aspects of algebra can be challenging even for highly educated adults.

The results are in keeping with the general tenants of cognitive load theory, whereby it is crucial to arrange the teaching environment in such a way that helps the learners to decrease the attentional and working memory demands of the presented algebraic material (Kirschner, 2002; Paas et al., 2003). The building of fluency in the execution of the component processes underlying the associated algebra, such as the sequence of problem-solving steps (Cooper & Sweller, 1987), and with basic arithmetic (Walczyk & Griffith-Ross, 2006) will reduce these attentional and working memory demands. Cooper and Sweller (1987) found that use of worked examples that provided information on the sequence of steps (problem-solving schema) involved in manipulating algebra equations improved student performance, likely attributable in part to reductions in working memory demands. Without explicit information on the sequence of problem-solving steps students often engage in attention demanding means–ends problem solving which is associated with increases in problem solving errors (Owen & Sweller, 1985). The basic idea is to commit relevant information, including problem solving schemas, to long-term memory such that their subsequent retrieval is not attention demanding (Sweller, 2016; Sweller et al., 2019).

Attention allocation is also necessary while forming new concepts in long-term memory (Cowan, 2014); effective learning occurs provided that the distractors do not interfere with the task (Fisher et al., 2014; Sweller, 2011). That is also in line with our experimental results; a tremendous decrease in spatial item recall (a source of attentional demand) in the presence of the more challenging algebraic task (expression evaluation). With this regard, one instructional approach to improve algebraic learning can be to eliminate or decrease the irrelevant information in the examples and explicitly point out the essential information or provide worked examples that includes this information for novice learners (Kirschner, 2002; Paas et al., 2003; Sweller et al., 2019). Another useful approach could be connecting algebraic topics with previously known mathematical areas, like arithmetic (for expression evaluation) and number line (for the coordinate plane), because decreasing the load on attention by using saved concepts is a strategy people use to form new long-term memories (Rhodes & Cowan, 2018).

Implications for Working Memory Theories

As mentioned above, we observed common attention-related effects of processing tasks on spatial load, independent of task nature (verbal vs spatial), and selectively on verbal load across all experiments. The former is consistent with a single attentional mechanism for all kinds of memory items (Cowan, 1999, 2019), but the less consistent influence on verbal load does not fit this mechanism. The absence of a clear effect of processing task on verbal load in three of the experiments (i.e., arithmetic BF = .17, geometry BF = .22, and mental rotation BF = .06) might be explained if a certain amount of verbal information can be rehearsed without relying heavily on attention (Morey et al., 2013; Vergauwe et al., 2014).

The notion of a rehearsal mechanism is of course consistent with a multicomponent model of working memory (e.g., Baddeley & Hitch, 1974; Baddeley & Logie, 1999) but it is not inconsistent with an embedded processes approach. Indeed, Cowan (1988; p. 165) said:

> The few items in short-term (active) storage can be maintained by mentally scanning or rehearsing the entire set. The vast amount of information in long-term storage cannot be scanned, but the retrievability can be improved by forming associations between items. This implies that there must be reliance on control processes more closely associated with phonetic characteristics (e.g., rote rehearsal) for short-term storage versus semantic characteristics (e.g., memory elaboration) for long-term storage, but the association would be imperfect.

Conversely, using attention for storage is an idea that in recent years has been a popular topic of investigation from the point of view of a modified version of the multicomponent model (Baddeley et al., 2019). The view of Barrouillet and Camos (2015) also includes both attention-based and articulation-based mnemonic mechanisms in working memory.

There is still a potential disagreement between the views regarding whether information in visual-spatial storage can be maintained without attention or not; the present finding of a ubiquitous effect of processing (spatial or otherwise) on spatial recall suggest that, in keeping with the embedded processes view (Cowan, 2019) and related views (Morey et al., 2013; Morey & Bieler, 2013; Vergauwe et al., 2010, 2014), attention is consistently needed for spatial storage.

Conclusion

Our study is the first to use, as far as we know, experimental manipulations to explore the role of domain-general attention and working memory mechanisms in algebra and to make causal inferences. The findings support the function of domain-general attention mechanisms and domain-specific verbal and spatial working memory processes in algebraic tasks throughout the five experiments. While verbal processing is prevalent in both expression evaluation and coordinate plane, the attentional needs are differential, depending on the task difficulty. The attentional and working memory demands of expression evaluation and placing ordered pairs in the coordinate plane, and presumably the demands of other algebraic tasks (Cooper & Sweller, 1987), have strong implications for instructional design; specifically, design features that reduce attentional demands during learning and that facilitate the formation of long-term memory representations of algebraic information should enhance student outcomes in algebra (Sweller et al., 2019).

Moreover, our main results contribute to the existing working memory research in several ways: (a) it demonstrates the role of the domain-general attentional system used in working memory regardless of task features (verbal or spatial), (b) provides further evidence regarding the dual maintenance mechanism of verbal information and differential attentional need of spatial and verbal items, and (c) postulates an explanation about the interaction between attentional requirements of the processing tasks and the verbal item maintenance mechanisms. Our findings also cultivate some critical questions for future research, including how domain-general attention contribute to the maintenance of working memory items, whether the verbal and spatial information is kept in storage while performing the algebraic tasks, and how the interplay between attention- and rehearsal- or articulation-based mechanisms of verbal mechanisms occurs.

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, *67*(1), 21–38. https://doi.org/10.1006/jecp.1997.2397

Allen, K., Giofrè, D., Higgins, S., & Adams, J. (2020). Working memory predictors of mathematics across the middle primary school years. *The British Journal of Educational Psychology*, *90*(3), 848–869. https://doi.org/10.1111/bjep.12339

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Academic Press.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory & Cognition*, *47*(4), 575–588. https://doi.org/10.3758/s13421-018-0878-5

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.005

Barbieri, C. A., & Booth, J. L. (2020). Mistakes on display: Incorrect examples refine equation solving and algebraic feature knowledge. *Applied Cognitive Psychology*, *34*(4), 862–878. https://doi.org/10.1002/acp.3663

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Psychology Press.

Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, *38*(3), 317–338. https://doi.org/10.1006/brcg.1998.1039

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992). Verbal recoding of visual stimuli impairs mental image transformations. *Memory & Cognition*, *20*(4), 449–455. https://doi.org/10.3758/BF03210929

Bürkner, P. C. (2019). *Bayesian item response modeling in R with brms and Stan*. arXiv. https://arxiv.org/abs/1905.09501

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 28. https://doi.org/10.18637/jss.v080.i01

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: The executive and phonological loops. *Frontiers in Human Neuroscience*, *8*, 900. https://doi.org/10.3389/fnhum.2014.00900

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Memory & Cognition*, *39*(2), 231–244. https://doi.org/10.3758/s13421-010-0011-x

Chen, Z., & Cowan, N. (2009). How verbal memory loads consume attention. *Memory & Cognition*, *37*(6), 829–836. https://doi.org/10.3758/MC.37.6.829

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*(4), 347–362. https://doi.org/10.1037/0022-0663.79.4.347

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cowan, N. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, *145*(8), 822–847. https://doi.org/10.1037/bul0000199

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Blume, C. L., Boone, A. P., & Saults, J. S. (2021). Developmental change in the nature of attention allocation in a dual task. *Developmental Psychology*, *57*(1), 33–46. https://doi.org/10.1037/dev0001134

Cowan, N., Donnell, K., & Saults, J. S. (2013). A list-length constraint on incidental item-to-item associations. *Psychonomic Bulletin & Review*, *20*(6), 1253–1258. https://doi.org/10.3758/s13423-013-0447-7

Cragg, L., Richardson, S., Hubber, P. J., Keeble, S., & Gilmore, C. (2017). When is working memory important for arithmetic? The impact of strategy and age. *PLoS ONE*, *12*(12), e0188693. https://doi.org/10.1371/journal.pone.0188693

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*(4), 924–942. https://doi.org/10.3102/0002831209349460

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, *37*(10), 1189–1199. https://doi.org/10.1016/S0028-3932(98)00159-6

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*(2), 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

DeStefano, D., & LeFevre, J. A. (2004). The role of working memory in mental arithmetic. The. *European Journal of Cognitive Psychology*, *16*(3), 353–386. https://doi.org/10.1080/09541440244000328

Dodson, C. S., Johnson, M. K., & Schooler, J. W. (1997). The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, *25*(2), 129–139. https://doi.org/10.3758/BF03201107

Doherty, J. M., Belletier, C., Rhodes, S., Jaroslawska, A., Barrouillet, P., Camos, V., Cowan, N., Naveh-Benjamin, M., & Logie, R. H. (2019). Dual-task costs in working memory: An adversarial collaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(9), 1529–1551. https://doi.org/10.1037/xlm0000668

Earnest, D. (2015). From number lines to graphs in the coordinate plane: Investigating problem solving across mathematical representations. *Cognition and Instruction*, *33*(1), 46–87. https://doi.org/10.1080/07370008.2014.994634

Ellis, D. M., Ball, B. H., Kimpton, N., & Brewer, G. A. (2020). The role of working memory capacity in analytic and multiply-constrained problem-solving in demanding situations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *73*(6), 920–928. https://doi.org/10.1177/1747021820909703

Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual environment, attention allocation, and learning in young children: When too much of a good thing may be bad. *Psychological Science*, *25*(7), 1362–1370. https://doi.org/10.1177/0956797614533801

Fuchs, L. S., Compton, D. L., Fuchs, D., Powell, S. R., Schumacher, R. F., Hamlett, C. L., Vernier, E., Namkung, J. M., & Vukovic, R. K. (2012). Contributions of domain-general cognitive resources and different forms of arithmetic development to pre-algebraic knowledge. *Developmental Psychology*, *48*(5), 1315–1326. https://doi.org/10.1037/a0027475

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition*, *28*(5), 774–782. https://doi.org/10.3758/BF03198412

Fyfe, E. R., Matthews, P. G., Amsel, E., McEldoon, K. L., & McNeil, N. M. (2018). Assessing formal knowledge of math equivalence among algebra and pre-algebra students. *Journal of Educational Psychology*, *110*(1), 87–101. https://doi.org/10.1037/edu0000208

Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school. *Educational Evaluation and Policy Analysis*, *22*(3), 241–254. https://doi.org/10.3102/01623737022003241

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*(4), 1343–1359. https://doi.org/10.1111/j.1467-8624.2007.01069.x

Geary, D. C., Hoard, M. K., Nugent, L., & Rouder, J. N. (2015). Individual differences in algebraic cognition: Relation to the approximate number and semantic memory systems. *Journal of Experimental Child Psychology*, *140*, 211–227. https://doi.org/10.1016/j.jecp.2015.07.010

Geary, D. C., Scofield, J. E., Hoard, M. K., & Nugent, L. (2021). Boys' advantage on the fractions number line is mediated by visuospatial attention: Evidence for a parietal-spatial contribution to number line learning. *Developmental Science*, *24*(3), e13063. https://doi.org/10.1111/desc.13063

Geary, D. C., & Widaman, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence*, *16*(1), 47–80. https://doi.org/10.1016/0160-2896(92)90025-M

Geary, D. C., Widaman, K. F., & Little, T. D. (1986). Cognitive addition and multiplication: Evidence for a single memory network. *Memory & Cognition*, *14*(6), 478–487. https://doi.org/10.3758/BF03202519

Gray, S., Green, S., Alt, M., Hogan, T. P., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, *92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004

Gronau, Q. F., Singmann, H., & Wagenmakers, E. J. (2020). bridgesampling: An R Package for Estimating Normalizing Constants. *Journal of Statistical Software*, *92*(10), 1–29. https://doi.org/10.18637/jss.v092.i10

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D. S., Forster, J. J., Wagenmakers, E.-J., & Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, *81*, 80–97. https://doi.org/10.1016/j.jmp.2017.09.005

Guitard, D., Saint-Aubin, J., Poirier, M., Miller, L. M., & Tolan, A. (2020). Forward and backward recall: Different visuospatial processes when you know what's coming. *Memory & Cognition*, *48*(1), 111–126. https://doi.org/10.3758/s13421-019-00966-w

Hadley, W., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Hadley, W., & Henry, L. (2019). *tidyr: Tidy messy data*. https://CRAN.R-project.org/package=tidyr

Hansen, M. (2014). Characteristics of schools successful in STEM: Evidence from two states' longitudinal data. *The Journal of Educational Research*, *107*(5), 374–391. https://doi.org/10.1080/00220671.2013.823364

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology*, *10*(3), 302–323. https://doi.org/10.1016/0010-0285(78)90002-6

Imbo, I., & Vandierendonck, A. (2007). Do multiplication and division strategies rely on executive and phonological working memory resources? *Memory & Cognition*, *35*(7), 1759–1771. https://doi.org/10.3758/BF03193508

Jeffreys, H. (1961). *Theory of probability*. University Press. https://doi.org/10.2307/2267987

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics* (Vol. 2101). National Academy Press.

Kirschner, P. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. https://doi.org/10.1016/S0959-4752(01)00014-7

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1–64. https://doi.org/10.3102/00346543060001001

Li, Y., & Geary, D. C. (2013). Developmental gains in visuospatial memory predict gains in mathematics achievement. *PLoS ONE*, *8*(7), e70160. https://doi.org/10.1371/journal.pone.0070160

Li, Y., & Geary, D. C. (2017). Children's visuospatial memory predicts mathematics achievement through early adolescence. *PLoS ONE*, *12*(2), e0172046. https://doi.org/10.1371/journal.pone.0172046

Martein, R., Kemps, E., & Vandierendonck, A. (1999). The role of working memory in a double span task. *Psychologica Belgica*, *39*(1), 15–27. https://doi.org/10.5334/pb.939

McNeil, N. M., Weinberg, A., Hattikudur, S., Stephens, A. C., Asquith, P., Knuth, E. J., & Alibali, M. W. (2010). A is for apple: Mnemonic symbols hinder the interpretation of algebraic expressions. *Journal of Educational Psychology*, *102*(3), 625–634. https://doi.org/10.1037/a0019105

Morey, C. C., & Bieler, M. (2013). Visual short-term memory always requires general attention. *Psychonomic Bulletin & Review*, *20*(1), 163–170. https://doi.org/10.3758/s13423-012-0313-z

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*(3), 324–348. https://doi.org/10.1016/j.jml.2013.04.004

National Mathematics Advisory Panel. (2008). *Foundations for success: Final report of the National Mathematics Advisory Panel*. Department of Education.

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology*, *77*(3), 272–284. https://doi.org/10.1037/0022-0663.77.3.272

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., Van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, *22*(2), 115–121. https://doi.org/10.1007/s10648-010-9133-8

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, *108*(4), 455–473. https://doi.org/10.1037/edu0000079

Pillay, H., Wilss, L., & Boulton-Lewis, G. (1998). Sequential development of algebra knowledge: A cognitive analysis. *Mathematics Education Research Journal*, *10*(2), 87–102. https://doi.org/10.1007/BF03217344

Pollack, C., Guerrero, S. L., & Star, J. R. (2016). Exploring mental representations for literal symbols using priming and comparison distance effects. *ZDM Mathematics Education*, *48*(3), 291–303. https://doi.org/10.1007/s11858-015-0745-9

Pollack, C., & Price, G. R. (2020). Mapping letters to numbers: Potential mechanisms of literal symbol processing. *Learning and Individual Differences*, *77*, 101809. https://doi.org/10.1016/j.lindif.2019.101809

Core Team, R. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rhodes, S., & Cowan, N. (2018). Attention in working memory: Attention is needed but it yearns to be free. *Annals of the New York Academy of Sciences*, *1424*(1), 52–63. https://doi.org/10.1111/nyas.13652

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Schönbrodt, F. D., & Stefan, A. M. (2018). BFDA: An R package for Bayes factor design analysis (Version 0.3) [Computer software]. https://github.com/nicebread/BFDA

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339. https://doi.org/10.1037/met0000061

Schnuerch, M., Heck, D. W., & Erdfelder, E. (2021). *Waldian t tests: Sequential Bayesian t tests with controlled error probabilities*. https://doi.org/10.31234/osf.io/x4ybm

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory & Cognition*, *22*(1), 1–13. https://doi.org/10.3758/BF03202756

Souza, A. S., & Oberauer, K. (2017). The contributions of visual and central attention to visual working memory. *Attention, Perception & Psychophysics*, *79*(7), 1897–1916. https://doi.org/10.3758/s13414-017-1357-y

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of

Educational Research, *81*(4), 453–492. https://doi.org/10.3102/0034654311423025

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sweller, J. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory & Cognition*, *5*(4), 360–367. https://doi.org/10.1016/j.jarmac.2015.12.002

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*(1), 59–89. https://doi.org/10.1207/s1532690xci0201_3

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Thalmann, I., Souza, A. S., & Oberauer, K. (2019). Revisiting the attentional demands of rehearsal in working-memory tasks. *Journal of Memory and Language*, *105*, 1–18. https://doi.org/10.1016/j.jml.2018.10.005

Trbovich, P. L., & LeFevre, J. A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*(5), 738–745. https://doi.org/10.3758/BF03196112

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239). Erlbaum.

Ünal, Z. E. (2021). *Working memory and algebra*. https://osf.io/g6mak/

Vandierendonck, A., De Vooght, G., & Van der Goten, K. (1998). Does random time interval generation interfere with working memory executive functions? *European Journal of Cognitive Psychology*, *10*(4), 413–442. https://doi.org/10.1080/713752284

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharsk, S., Ly, A., Marsman, M., Matzke, D., Komarlu Narendra Gupta, A. R., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E. J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*, 813–826. https://doi.org/10.31234/osf.io/yqxfr

Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science*, *21*(3), 384–390. https://doi.org/10.1177/0956797610361340

Vergauwe, E., Camos, V., & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1072–1095. https://doi.org/10.1037/a0035779

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250–270. https://doi.org/10.1037/0033-2909.117.2.250

Walczyk, J. J., & Griffith-Ross, D. A. (2006). Time restriction and the linkage between subcomponent efficiency and algebraic inequality success. *Journal of Educational Psychology*, *98*(3), 617–627. https://doi.org/10.1037/0022-0663.98.3.617

Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064. https://doi.org/10.3758/s13423-012-0295-x

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 6–10. https://doi.org/10.3758/BF03202594

Appendix A

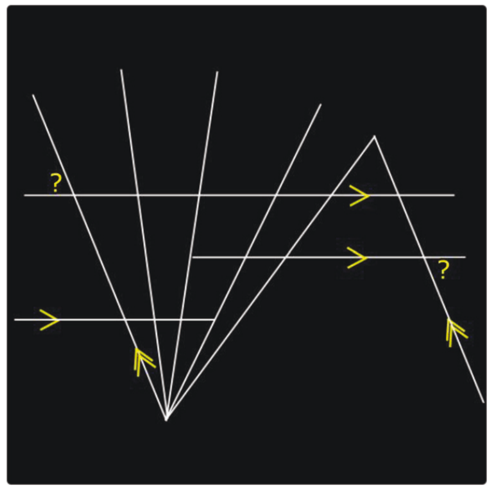
Expression Evaluation Questions

$x = 6$	$y = 9$	$7x + y = 52$
$x = 7$	$y = 8$	$8y + x = 71$
$y = 18$	$x = 9$	$6x + y = 62$
$y = 4$	$x = 15$	$9y + x = 51$
$x = 7$	$y = 9$	$8x + y = 66$
$x = 9$	$y = 12$	$6y - x = 64$
$y = 8$	$x = 13$	$7x - y = 83$
$y = 12$	$x = 7$	$8y - x = 88$
$x = 8$	$y = 4$	$4x - y = 28$
$y = 6$	$x = 5$	$4y - x = 19$
$x = 7$	$y = 8$	$7x + y = 58$
$x = 9$	$y = 8$	$8y + x = 73$
$y = 15$	$x = 2$	$9x + y = 32$
$y = 8$	$x = 3$	$6y + x = 51$
$x = 3$	$y = 17$	$9x + y = 45$
$x = 8$	$y = 14$	$6y - x = 76$
$y = 7$	$x = 12$	$7x - y = 77$
$y = 13$	$x = 6$	$7y - x = 75$
$x = 8$	$y = 9$	$4x - y = 24$
$y = 6$	$x = 8$	$6y - x = 28$
$x = 8$	$y = 7$	$7x + y = 64$
$x = 9$	$y = 6$	$9y + x = 63$
$y = 9$	$x = 7$	$6x + y = 50$
$y = 3$	$x = 18$	$8y + x = 42$
$x = 4$	$y = 16$	$9x + y = 53$
$x = 9$	$y = 13$	$6y - x = 70$
$y = 6$	$x = 12$	$7x - y = 78$
$y = 12$	$x = 8$	$8y - x = 78$
$x = 9$	$y = 7$	$4x - y = 29$
$y = 8$	$x = 5$	$4y - x = 27$

(*Appendices continue*)

Appendix B

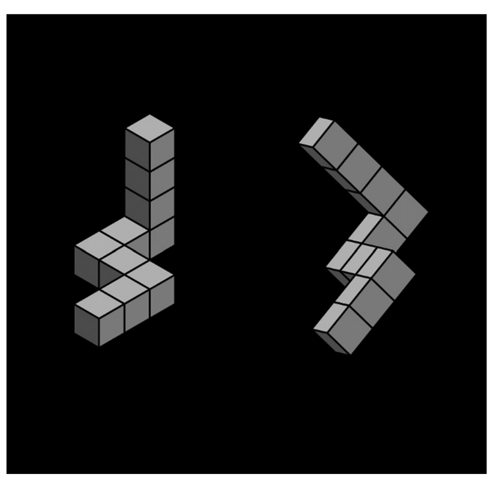
An Example of Geometry Questions



Note. See the online article for the color version of this figure.

Appendix C

An Example of Mental Rotation Figures



Received March 5, 2021
Revision received October 19, 2021
Accepted November 26, 2021 ■