Empirical Research

Spacing Out! Manipulating Spatial Features in Mathematical Expressions Affects Performance

Avery Harrison*[a] [iD] , Hannah Smith[a] [iD] , Taylyn Hulse[a] [iD] , Erin R. Ottmar[a] [iD]

[a] Department of Social Science & Policy Studies, Worcester Polytechnic Institute, Worcester, MA, USA.

Abstract

The current study explores the effects of physical spacing within mathematical expressions on student performance. A total of 2,152 students in 5th-12th grade were randomly assigned to one of four conditions within an online problem set, with terms in algebraic expressions spaced 1) neutrally, with no spaces in the expression, 2) congruent with the order of precedence through grouping terms, 3) incongruent with the order of precedence, or 4) mixed, a combination of the previous conditions. Results show that students who viewed incongruent problems made more errors and had to solve more problems to complete the assignment than those who viewed congruent or neutrally spaced problems. Additionally, students who viewed problems with mixed spacing had to solve more problems to complete the assignment than students who viewed congruent problems. These findings suggest that viewing expressions with spacing that is incongruent with the order of precedence presents challenges for students. Overall, these results replicate prior research in perceptual learning in a natural homework environment and support the claim that physical spacing between terms does influence student performance on order of precedence problems.

Keywords: perceptual learning, spatial proximity, mathematical cognition, mathematical operations

Formal mathematics is a commonly used example of how humans make sense of abstract symbolic reasoning (Anderson, 2007; Goldstone et al., 2017). However, learning mathematics is difficult for many students, in part because of the requirement to learn and execute abstract rules as they apply to mathematical notation. Being able to interpret symbolic notation and compute simple calculations efficiently and accurately is critical for solving more complex mathematics problems, notably algebra. For example, the order of precedence stipulates how to simplify an expression or equation, including the order in which computations can be carried out. Such abstract rules require students to learn seemingly arbitrary conceptual processes and appropriately apply those rules when reasoning about mathematics.

Beyond being abstract and requiring conceptual knowledge, reasoning about mathematics is also inherently perceptual (Marghetis et al., 2016) with ample evidence suggesting that mathematical processing and understanding is influenced by the visual presentation of mathematical notation (McNeil & Alibali, 2004, 2005). Perceptual learning has been suggested to be a mechanism that adapts perception and directs attention to relevant information in the environment (Gibson, 1969), supporting high level cognition. For instance, spatial

proximity, a Gestalt law that posits that individuals perceive objects in close proximity to be a group, has been shown to bias mathematical reasoning. This phenomenon in mathematics supports the notion that people rely on perceptual cues to process symbolic notations and are heavily influenced by spatial properties of notation (Goldstone et al., 2017; Wagemans et al., 2012). Regardless of conceptual knowledge, the tendency to use perceptual cues and groupings in mathematics notation is somewhat automatic and has implications for the ways in which individuals interpret, compute, and produce mathematics notation.

Although subtle visual manipulations are irrelevant to the mathematical meaning of notation, visual manipulations of notation can lead to attentional biases and create perceptual groupings among terms and operands. For instance, terms and operands spaced in close proximity within a mathematical notation tend to be seen as a group, such as viewing "4 x 6+3" and wrongly grouping "6+3" together based on the spatial proximity of those terms (Jiang, Cooper, & Alibali, 2014; Kirshner, 1989; Landy & Goldstone, 2007b, 2010). Additionally, while novice learners often solve order of precedence problems based on memorized rules, experts have been shown to rely on perceptual cues when solving complex equations (Braithwaite et al., 2016; Rumelhart et al., 1986), providing evidence that there may be a shift at some point in experience or procedural fluency from attending to abstract rules of formal mathematics to attending to perceptual cues in formal mathematical notation.

A large body of research has demonstrated that the physical spacing between terms and operands within equations and expressions contributes to students' perceptions of *how* they are able, within the rules of mathematics, to interpret meaning and perform computations (e.g., Jiang et al., 2014; Landy & Goldstone, 2007a, 2007b, 2010; Rivera & Garrigan, 2016). Consequently, spacing in mathematical expressions has been found to impact performance on equation-solving. For instance, Landy and Goldstone (2007b, 2010) manipulated whether the spacing of terms in expressions was congruent (multiplications spaced closer than additions) or incongruent (additions spaced closer than multiplications) with the order of precedence. They found that participants made more errors and were more likely to perform addition before multiplication in the incongruent spacing condition. For example, in the case of 7+1 * 4, people often first combine the 7+1 to make 8 and then multiply by 4 to get 32, instead of properly multiplying 1 by 4 and then adding 7. These results suggest that mathematical reasoning is at least somewhat perceptually driven through low-level visual and attentional factors. Landy and Goldstone (2007b, 2010) posit that this effect occurs because spacing cues bias individuals to perform specific operations, even if those cues are mathematically invalid.

Similarly, Jiang and colleagues (2013) found that when participants viewed operand spacing in expressions which created perceptual groupings incongruent with the order of precedence, participants tended to make target errors reflective of incorrectly grouping a set of terms. Rivera and Garrigan (2016) extended this work by replicating the effect of incongruent spacing on order of precedence errors found by Landy and Goldstone (2010), providing further support for the effects of perceptual grouping on mental arithmetic, even in the case of evaluating simple expressions. This work provides evidence that when perceptual grouping is incongruent with operator precedence, the likelihood of order of precedence errors in mental arithmetic increases. More broadly, this research shows that individuals use perceptual spacing to interpret and reason about mathematics and may have a difficult time ignoring perceptual cues even if they are incongruent with mathematical rules.

Gómez, Benavides-Varela, Picciano, Semenza, and Dartnell (2014) extended this work with a sample of 5th-8th Chilean and Italian students and found that the spacing effects seem to emerge in younger students

as well. Braithwaite et al. (2016) also explored the effect of physical spacing outside of a laboratory setting among even younger primary-school children (equivalent to U.S. grade levels 2-6) in the Netherlands and found higher error rates for individuals who viewed problems which had spacing incongruent with the order of precedence. They also found that this effect of spacing increased with grade level, further suggesting that there is an increased reliance on perceptual grouping with age and experience with arithmetic. However, aside from the work of Gómez, Benavides-Varela, et al. (2014) and Braithwaite et al. (2016), less is known about whether or how perceptual grouping, influenced by physical spacing within mathematical expressions, impacts student behavior in typical school settings and varies across grade levels.

Overall, this body of literature shows that the perceptual grouping of mathematical terms in an expression or equation influences both novices and experts during problem solving. Specifically, when terms are spatially organized in groups that mirror the order of precedence, students are more likely to have higher performance (Landy & Goldstone, 2007a) and more accurate interpretations (Jiang et al., 2014; Landy & Goldstone, 2010). Conversely, when terms are grouped in ways that are incongruent with the order of precedence, students are more likely to take more time to solve (Gómez, Bossi, & Dartnell, 2014) and make more errors (Jiang et al., 2013; Landy & Goldstone, 2007b, 2010; Rivera & Garrigan, 2016). Such research provides evidence of the influence of perceptual learning on mathematical problem solving, which could play a key role in student learning. To further this area of research, the current study aims to replicate and extend prior research by exploring the effects of spacing on student performance on order of operations problems with upper elementary through high school algebra learners in a natural homework setting.

The present study asked 5-12th grade students to simplify order of precedence expressions in ASSISTments, an online tutoring system (Heffernan & Heffernan, 2014). Students were randomized into one of four experimental conditions, which manipulated the physical spacing between numbers and terms within mathematical expressions to be either *neutral*, *congruent* or *incongruent* with the order of precedence, or a *mixed* combination of spacing across problems. We then examined whether there were differences in content mastery speeds (the total number of problems that students had attempted by the time that they correctly answered three problems in a row) based on spacing conditions.

This study extends prior research on perceptual learning in four key ways. First, the majority of studies examining the effects of physical spacing between mathematical terms have been conducted with undergraduate students in controlled laboratory settings rather than with school-aged children in authentic classroom and learning contexts. Second, the study is conducted through an online homework assignment assigned to students by their teachers using the ASSISTments platform (Heffernan & Heffernan, 2014), rather than administered by a researcher in a laboratory setting. Third, this study examines the effect of a mixed condition, where students are exposed to each of the experimental spacing conditions. Lastly, while many studies on perceptual learning have used error rates as the learning outcome, the current study uses mastery speeds, a measure of the number of problems attempted to master the material presented in the assignment as the dependent measures. From this extension of related research, this project aims to contribute a richer understanding of how perceptual grouping, from physical spacing in mathematical expressions, affects students' behavior in authentic learning contexts.

The Present Study

To extend prior research on perceptual learning as it pertains to mathematics performance, we present a randomized controlled trial with upper elementary, middle, and high school students in ASSISTments, an online tutoring system. This study is designed to explore the impact of physical spacing between terms on students' mastery speeds when solving a series of order of operations problems. Specifically, we investigate our hypotheses by posing the following questions:

1. *Does spacing impact assignment mastery speed?* We hypothesize that students who view congruent or neutrally spaced problems will have quicker mastery speeds (attempting fewer problems before correctly answering three problems in a row) compared to students who view incongruent spacing or mixed spacing problems.

2. If there are differences in assignment mastery speed based on condition, *does student prior performance moderate the relationship between condition and mastery speed?* We explore possible interactions between condition and prior performance to see if different levels of prior performance heighten or mitigate the effect of any spacing condition(s) on mastery speeds.

3. If there are differences in assignment mastery speed based on condition, *does grade level moderate the relationship between condition and mastery speed?* We explore possible interactions between condition and grade to see if the effect of any spacing condition(s) on mastery speed varies by grade level.

Method

Study Context

Data for this study was collected from 2015-2019 in ASSISTments, an online tutoring system that features free content for K-12 students with a primary focus on mathematics (Heffernan & Heffernan, 2014). In addition to providing a technology tool for teachers to assign content and homework to students, ASSISTments also provides researchers with an experimental platform where independent researchers can create their own randomized controlled trials to be used by teachers and students. The de-identified data used in this analysis is available on Open Science Framework (Harrison, Smith, Hulse, & Ottmar, 2020b). Additionally, the original data report from ASSISTments is available for further reference (Harrison, Smith, Hulse, & Ottmar, 2020a) in the Supplementary Materials.

This randomized controlled trial was created by the authors and deployed as an available Skill Builder problem set covering order of operations content (targeting 7th grade) within ASSISTments. "Skill Builders" are optional problem sets that teachers can assign to provide students with fluency practice on topics commonly featured on standardized mathematics tests. Skill Builders map onto content areas from the Common Core State Standards and present problems from a given content area in a randomized order. These problem sets are designed to challenge a student in a mathematics topic until that student achieves content mastery.

Under default settings, students must consecutively answer three problems in a row correctly to achieve mastery status for the Skill Builder assignment. If a student answers a problem incorrectly, the problem count restarts and they continue to receive problems until they correctly answer three problems in a row. Therefore,

in this context, a slower mastery speed (solving more problems in order to get three problems correct in a row) is an indicator of higher error and lower mathematics performance on a Skill Builder assignment. Mastery speed has been used as an outcome measure of student performance in previous ASSISTments studies (e.g., Botelho et al., 2015).

Participants

The final sample included in the analyses were 2,152 students (48.0% male, 35.2% female, 16.9% unknown) who completed more than three problems in the Skill Builder problem set and completed the assignment by achieving mastery. Participants were 5th-12th grade students assigned to complete the given problem set by their classroom teacher. The 2,152 students included in the final sample from this study came from 199 classes taught by 115 teachers from 83 schools in 64 districts from 16 states. The students were distributed across several grade levels, with a majority of students in middle school classrooms (0.6% fifth, 11.7% sixth, 30.4% seventh, 9.8% eighth, 18.5% ninth, 1.0% tenth, 0.1% eleventh, and 1.1% of reported cases in twelfth grade; with the remaining 26.9% of cases missing grade level information).

Many more students initially opened the problem set but were dropped from this study for the following reasons. A total of 6,238 students opened the problem set, however, 4,053 students were excluded due to assignment completion within three problems or stopping the assignment within the first three problems, thus never seeing an experimental condition. Additionally, a small subset of participants was also excluded due to having an unknown mastery status for the problem set (n = 33).

Experimental Conditions and Procedure

When students opened the problem set, they were first exposed to three neutrally spaced expressions to solve (Figure 1). After completing the three neutrally-spaced problems, students were randomly assigned to one of four spacing conditions: 1) *neutral* (n = 574), with no spaces in the expression, 2) *congruent* (n = 555), with spacing which follows the order of precedence, 3) *incongruent* (n = 493; see Figure 2), with spacing which does not follow the order of precedence, or 4) *mixed* (n = 530), a combination of the previous conditions. Once assigned to a condition, students were presented with additional problems to solve. The problems in each condition were identical in structure but varied in the physical spacing of terms within each expression. The first several problems for each condition are shown in Table 1.

Assignment: Order of Operations no exponents 7.NS.A.3. EX

Problem ID: *PRA3RVW*   *Comment on this problem*

6*3+4*4

Type your answer below (mathematical expression):

[Submit Answer]   100% ⓘ   [Show hint]

Figure 1. Problem with neutral spacing as shown in first three assignment problems.
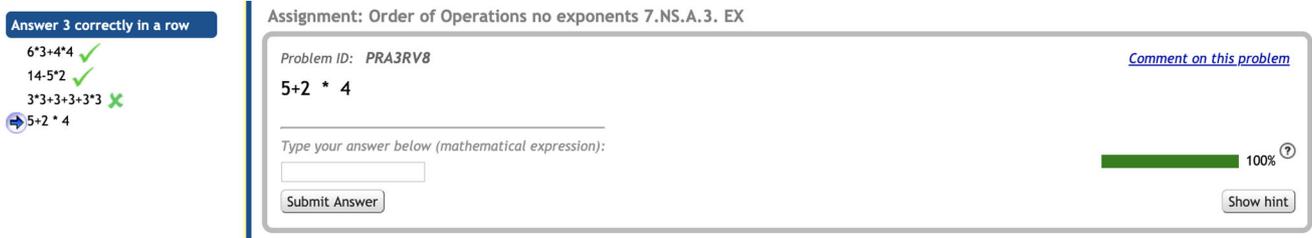
Figure 2. Example assignment screen for a participant in the incongruent spacing condition.

Table 1

The First Eight Problems Assigned by Condition

	Tutorial: All Participants			
1.	6*3+4*4			
2.	14-5*2			
3.	3*3+3+3*3			
	Neutral	**Congruent**	**Incongruent**	**Mixed**
4.	5+2*4	5 + 2*4	5+2 * 4	5 + 2*4
5.	7*2+8*5	7*2 + 8*5	7 * 2+8 * 5	7 * 2+8 * 5
6.	4*3+2	4*3 + 2	4 * 3+2	4*3+2
7.	4*(2+5)+12-2*3	4*(2 + 5) + 12 - 2*3	4 * (2+5)+12-2 * 3	4*(2 + 5) + 12 - 2*3
8.	5+3*2	5 + 3*2	5+3 * 2	5+3 * 2

Most students continued to solve problems until they achieved mastery (answering three consecutive problems correctly on the first try). However, if a student answered a problem incorrectly, they could not move on to the next problem until typing in the correct answer. To support students as they moved through the assignment, one hint restating the order of operations was available to click on at the beginning of each problem (Figure 3). If students elected to see the hint, they were then immediately able to click "Show Answer" which would display the correct answer to type as the solution. Importantly, if students opted to view the hint, the problem was marked as incorrect and did not count towards the three mastery problems required to complete the assignment. Additionally, students could opt to stop the assignment at any time without completion. However, for students that did achieve mastery status, ASSISTments automatically closed the assignment and marked the status as completed.
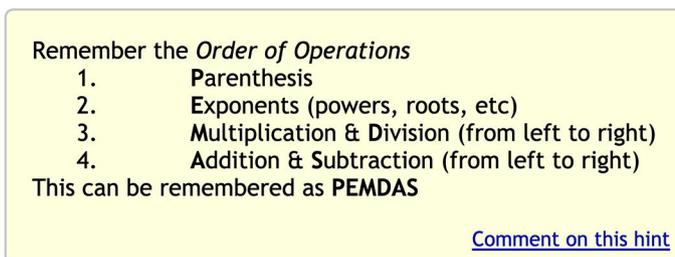
Figure 3. Hint available to participants on each problem in the assignment.

The study remained open as an active Skill Builder for the order of operations standard without exponents (Common Core Standard 7.NS.A.3. EX) that teachers could easily assign to their students at any time for

three years. At the end of the three years, the data from the study was aggregated using the ASSISTments Assessment of Learning Infrastructure (ALI) report that was automatically generated by the ASSISTments team for external researchers and provides aggregated data files at various levels of granularity such as student-level and problem-level (Ostrow et al., 2016). All variables of interest in this study were extracted from this report and are described in more detail below.

Measures

Prior to data analysis, the following measures for analyses were defined and extracted from the ASSISTments report as necessary for analysis.

Prior Mathematics Performance

As an estimated measure of prior mathematics performance, ASSISTments calculates a prior proportion correct value (from 0-1). This value represents the proportion of all previous ASSISTments problems completed from other assignments that each student answered correctly prior to the current experiment. However, the type of content may have varied and some participants may have had extensive experience with ASSISTments over years whereas others might have been first- or second-time users. Although participants varied in previous exposure and practice with ASSISTments, this value serves as a proxy for prior mathematics performance and has been used in studies that were deployed using the ASSISTments platform (e.g., Walkington et al., 2019). The distribution of prior performance scores was bimodal only due to a small subset (3.5%) of students who had demonstrated perfect prior performance in ASSISTments. This value was used as a continuous covariate for prior performance in analyses.

Grade Level

The ASSISTments ALI report also provided an ordinal value representing the reported grade level of each participant by the classroom teacher. With values ranging from 5-12, grade level was treated as a continuous variable in all analyses.

Approach to Analysis

After preprocessing the data, descriptive statistics were calculated in SPSS to determine means and variability for each variable and relations between each construct. Next, we conducted a one-way analysis of variance (ANOVA) with condition (neutral, congruent, incongruent, and mixed) as the independent variable and mastery speed as the dependent measure. We also conducted post hoc tests with Bonferroni correction to examine where there were significant differences in average mastery speed between conditions.

In addition to the ANOVA, we examined the impact of condition, above and beyond prior performance and grade level. To determine whether or not multilevel analysis would be appropriate, we calculated the intraclass correlation coefficient (ICC) from an unconditional 2-level hierarchical linear model (HLM; Model 1). An unconditional HLM model predicting mastery speed suggested that approximately 10% of the variance in mastery speed was attributable to differences at the class level. As this value exceeds the 7% variance threshold to suggest that using HLM would be appropriate (Lee, 2000; Niehaus et al., 2014), we chose to use HLM for all analyses to account for the nesting of students in classes.

PsychOpen GOLD

Next, four two-level HLMs were conducted to explore our research questions. Model 2 estimates how the covariates, grade level and prior performance, impact participants' assignment mastery speed while accounting for any nested effects between the student and class levels. Model 3 includes the three condition variables (neutral, congruent, and mixed, with incongruent as the reference group) and estimates how the physical spacing between terms impacts participants' assignment mastery speed (compared to the incongruent condition) while accounting for any nested effects between the student and class levels.

Model 3 in HLM has the following form:

$$\text{MASTERYSPEED}_{ij} = \gamma_{00} + \gamma_{10}(\text{CLASSGRADE}_{ij}) + \gamma_{20}(\text{NEUTRAL}_{ij}) + \gamma_{30}(\text{CONGRUENT}_{ij}) + \gamma_{40}(\text{MIXED}_{ij}) + \gamma_{50}(\text{PRIORPERC}_{ij}) + \mu 0_j + r_{ij}$$

where i is Students 1 through n, and j is Class 1 through n.

Interaction terms were created and added to the hierarchical linear model to examine interactions between *prior performance and condition* as well as *grade level and condition* as predictors of mastery speed, controlling for prior performance and grade level as covariates. Model 4 presents results for the second research question, exploring whether an interaction between prior performance and spacing condition predicted mastery speed. Lastly, Model 5 presents results for the third research question, which explores whether an interaction between grade level and spacing condition predicted mastery speed.

Results

Descriptive Statistics

Overall, all students completed the assignment by eventually achieving mastery status (answering three problems correctly in a row) at some point in the assignment (M = 6.38 problems, SD = 3.24 problems). While working on the problem set, 33.6% of participants used the available hint at least once. See Table 2, below, for details on students' prior performance and average mastery speed by grade level. Prior performance scores indicated that, on average, students had correctly answered 70% of previously attempted problems in ASSISTments prior to the beginning of this study (M = .70, SD = .14).

Table 2

Descriptive Statistics on Student Performance by Grade Level

Population	n	Average Prior Performance (*SD*)	Average Mastery Speed (*SD*)
Overall	2,152	.70 (.14)	6.38 (3.24)
5th Grade	13	.86 (.15)	7.23 (5.00)
6th Grade	251	.71 (.13)	6.34 (2.96)
7th Grade	654	.67 (.14)	6.39 (3.13)
8th Grade	210	.76 (.15)	6.25 (3.72)
9th Grade	399	.71 (.13)	5.90 (2.44)
10th Grade	21	.77 (.11)	6.57 (3.16)
11th Grade	2	.72 (.09)	5.50 (0.71)
12th Grade	24	.73 (.08)	6.79 (3.90)

Note. Grade level was not reported for n = 578 participants.

Next, we conducted a preliminary one-way ANOVA to examine differences in average mastery speeds by condition. Results indicate that there were statistically significant overall differences between groups in mastery speed, $F(3, 2148) = 10.33$, $p < .01$. Post hoc tests using Bonferroni correction to account for multiple comparisons revealed that, on average, students in the congruent condition (M = 5.94 problems, SD = 2.29 problems) mastered the assignment in significantly fewer problems than in the incongruent condition (M = 6.95 problems, SD = 3.58 problems, Cohen's d = 0.34) and the mixed condition (M = 6.59 problems, SD = 3.93 problems, Cohen's d = 0.20); see Figure 4. Students in the neutral condition (M = 6.14 problems, SD = 2.91 problems) also completed the problem set in significantly fewer problems than students in the incongruent condition (Cohen's d = 0.25). There were no differences in mastery speed between the mixed and incongruent condition ($p > .10$). These results prompted further exploration into examining the impact of spacing condition on assignment mastery speeds, accounting for grade level, prior performance, and the nesting of students in classrooms.
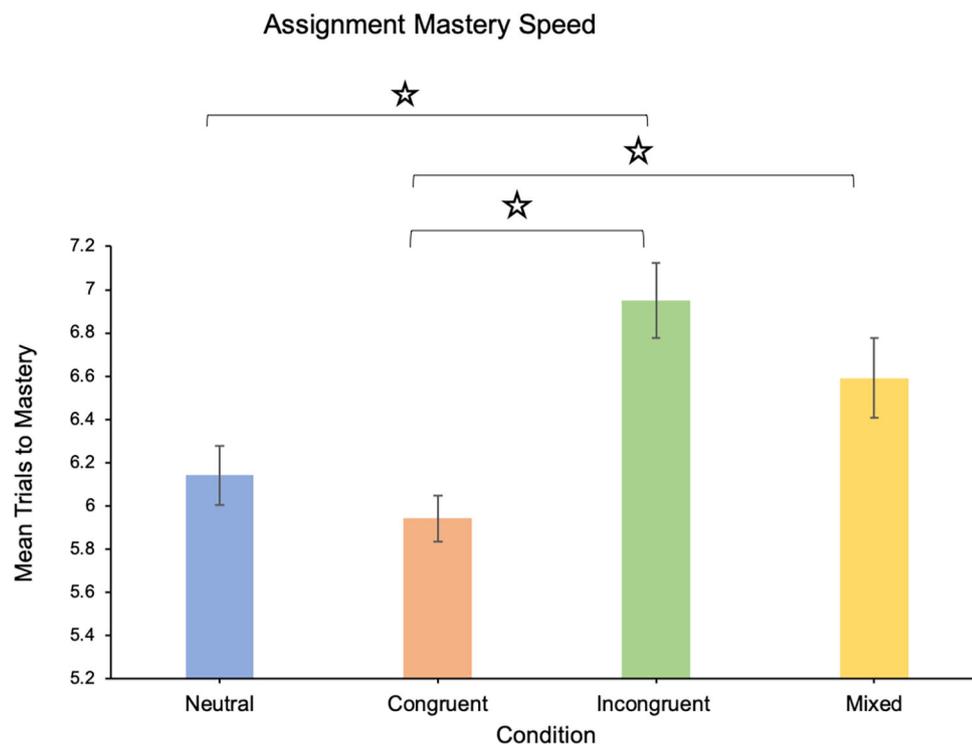

Figure 4. Mean number of trials required for mastery as a function of spacing condition with error bars reporting one standard error of the mean on each side.

Note. The star symbol denotes statistically significant ($p < .01$) between group differences.

Hierarchical Linear Models Examining the Impact of Spacing Condition on Mastery Speed

Table 3 (below) displays the results of the four two-level hierarchical linear models for these analyses. The unconditional model (Model 1) predicting mastery speed had an ICC of .097, indicating that 9.7% of the variance in assignment mastery speed is due to class level differences. The percentage of variance explained for each model is derived from the variance components of the model directly preceding it, explaining the variance accounted for above and beyond the previous model.

PsychOpen GOLD

Table 3

Hierarchical Linear Models Show the Effect of Condition, Prior Performance, and Grade on Mastery Speed

Parameter	Model 1		Model 2		Model 3		Model 4		Model 5	
	β	SE	β	SE	β	SE	β	SE	β	SE
Intercept	6.42**	0.11	6.23**	0.11	6.79**	0.22	7.67**	1.04	8.04**	1.09
Level-1 (Student)										
Grade Level			-0.01	0.09	-0.02	0.09	-0.02	0.09	-0.22	0.20
Prior Performance			-2.31**	0.69	-2.32**	0.70	-3.70*	1.61	-2.33**	0.70
Neutral					-0.78**	0.24	-2.21	1.77	-3.62*	1.72
Congruent					-0.92**	0.26	-2.68*	1.35	-2.43	2.15
Mixed					-0.46†	0.24	-0.82	1.70	-2.33	1.54
Neutral x Prior Performance							2.02	2.37		
Congruent x Prior Performance							2.51	1.77		
Mixed x Prior Performance							0.52	2.27		
Neutral x Grade									0.37†	0.21
Congruent x Grade									0.20	0.25
Mixed x Grade									0.25	0.18
Level-2 (Teacher)										
Variance Components										
Student Level	9.52		8.90		8.80		8.79		8.78	
Teacher Level	1.02		0.43		0.42		0.43		0.43	
Total Variance					10.54					
Level 1					0.90					
ICC Level 2					0.10					
% Explained at Student Level			0.07		0.01		0.00		0.00	
% Explained at Classroom Level			0.58		0.02		-0.02		0.00	

$^{\dagger}p < .10.$ $^{*}p < .05.$ $^{**}p < .01.$

Model 2 shows the influence of students' grade level and prior performance in ASSISTments on assignment mastery speed. Prior performance was a significant predictor of mastery speed, where students with higher prior performance on ASSISTments problem sets had lower mastery speeds (β = -2.31, *p* < .05). Grade level did not significantly predict mastery speed (*p* > .05). The addition of these two variables explained 7% of the child level variance and 58% of the class level variance in mastery speed.

Model 3: Research Question 1: Does Spacing Influence Assignment Mastery Speed?

A 2-level HLM model (Model 3) was conducted to examine the impact of condition on mastery speed, controlling for grade and prior performance. The incongruent spacing condition was treated as the reference group for the hierarchical linear models since the ANOVA indicated that there were significant differences between the incongruent spacing condition and two other groups. Results were consistent with the ANOVA; there was a significant effect of two conditions on assignment mastery speeds. The analysis revealed that students in the congruent condition (β = -0.92, *p* < .01) and the neutral condition (β = -0.78, *p* < .01) mastered the assignment in significantly fewer problems than in the incongruent condition. Specifically, students in the congruent condition completed the assignment, on average, in 0.92 problems faster than students in the incongruent condition. Similarly, students in the neutral condition completed the assignment, on average, in 0.78 problems quicker than students in the incongruent condition. While both congruent and neutral spacing conditions significantly

predict assignment mastery speed, the effect is larger for the congruent spacing condition than the neutral spacing condition. Although there was a trend, there were no significant differences between the mixed condition and the incongruent condition (β = -0.46, p < .07). Higher prior performance was related to faster mastery speeds (β = -2.32, p < .01). Grade level was not significant in predicting mastery speed.

Model 4: Research Question 2: Does Prior Performance Moderate the Effect of Condition on Assignment Mastery Speed?

Next, we tested whether there was an interaction effect between students' prior performance and condition to understand if the effect of spacing condition was moderated by prior performance. Model 4 presents all Level-1 variables in addition to the prior performance interaction terms within Level-1. The interactions between prior performance and each spacing condition (neutral, congruent, mixed) were not significant predictors of assignment mastery speed (β = 2.02, p = .39; β = 2.51, p = .16; β = 0.52, p = .82), which means only the main effects of prior performance and spacing condition should be interpreted. This result suggests that students' prior performance in ASSISTments does not moderate the relationship between spacing condition and assignment mastery speed.

Model 5: Research Question 3: Does Grade Moderate the Effect of Condition on Assignment Mastery Speed?

Lastly, we tested whether there were interaction effects of grade level × condition on assignment mastery speed. Model 5 presents all Level-1 variables in addition to the grade level interaction terms within Level-1. The interaction between grade level and the neutral spacing condition was not significant but was trending towards significance (β = 0.37, p = .08). The other interactions between grade level and each spacing condition (congruent, mixed) were not significant predictors of assignment mastery speed (β = 0.20, p = .43; β = 0.25, p = .18). This finding suggests that students' grade level does not moderate the relationship between spacing condition and assignment mastery speed.

Discussion

The goal of this study was to explore whether manipulating the physical spacing between mathematical symbols would impact students' assignment mastery speed on order of operations problems. In addition to replicating the difficulty that algebra learners experience with incongruent spacing in order of operations problems, we were particularly interested in examining whether spacing effects exist in both younger and older students in authentic homework environments such as an online tutoring system. Two main findings emerged from this study: 1) students in the incongruent condition had slower mastery speeds (solving more problems to achieve mastery) than students in the congruent or neutral conditions, and 2) there were no significant interactions between grade level and condition, or prior performance and condition, on mastery speeds. Together, these results suggest that viewing incongruent spacing within mathematical expressions led to more errors and lower performance for most students, regardless of age or prior performance, compared to those who viewed problems with congruent or neutral spacing between terms.

Physical Spacing in Math Expressions Affects Student Performance

We predicted that viewing congruent or neutral spacing within problems would lead to faster mastery speeds compared to viewing problems with incongruent or mixed spacing. The results mostly supported this hypothesis; students who viewed problems with congruent or neutral spacing tended to master the assignment in significantly fewer problems than students who viewed problems with spacing that was incongruent with the order of precedence. However, there were no significant differences in mastery speed between the neutral and mixed condition.

One explanation for why congruent spacing may lead to greater performance over incongruent spacing is that visually modifying the physical spacing of terms may bias people to naturally group proximal terms into grouped objects (Wertheimer, 1950). Building on this visual spacing bias, one could argue that perceptually grouping terms to be congruent with the order of precedence could be more advantageous for students by providing perceptual cues that direct attention towards higher precedence operations in expressions, as if providing visual scaffolding for the order of precedence within expressions. If this is the case, then it could be hypothesized that students who saw the congruent grouping should also be more likely to have faster mastery speeds than students in the neutral spacing condition. However, there were no clear advantages of using congruent as opposed to neutral spacing. This finding was aligned with those of Landy and Goldstone (2010) who also found that operation error rates did not differ between congruent and neutral spacing conditions.

The finding that viewing congruent and neutral spacing led to higher performance than incongruent spacing is consistent with prior studies (e.g., Gómez, Bossi, & Dartnell, 2014; Jiang et al., 2013, 2014; Landy & Goldstone, 2010; Rivera & Garrigan, 2016). An interpretation of these results is that while the visual structure of mathematical notation creates perceptual groupings that cue interpretation and computation biases, this effect is stronger when those groupings are incongruent with mathematical rules, knowledge, and the order of precedence. It is possible that the difference in mastery speed by condition is due to a reliance on multiple perceptual cues that individuals use when solving order of operations problems. Further, these cues may work in a hierarchical structure where physical spacing acts as a first-order perceptual cue and operands act as second-order cues to interpret and act on mathematical notation. As a result, when presented with incongruent spacing, students may (incorrectly) attend to and rely more on perceptual groups when simplifying an expression than when presented with congruent or neutral spacing.

Other work has suggested that spacing is used as an action-guiding cue; incongruent spacing elicits errors while congruent and neutral spacing in mathematical notation helps facilitate improved performance. Consequently, viewing congruent spacing in expressions may not be significantly more helpful than viewing neutral spacing because the perceptual cues from physical spacing would be redundant to cues from operands. This notion is supported by findings that individuals attend to multiplication operands quicker than addition operands and treat narrow spacing between terms similarly to multiplication operands (Landy et al., 2008). For instance, in the expression "4+2*3", the multiplication operand acts as a cue to group the "2" and "3" when the physical spacing is neutral. If the expression was presented with physical spacing congruent with the order of precedence, "4 + 2*3", the physical spacing would only reinforce the grouping between the "2" and "3" but is not necessary. Conversely, physical spacing would be more of a perceptual cue if the expression was presented as "4+2 * 3". Based on the results from this study and examples from the body of literature on this work demonstrating the influence of visual properties on performance with mathematics notation, perhaps the

physical spacing within mathematical notation is a higher-order perceptual cue than operands which is why viewing incongruent spacing may be much more challenging for students.

Prior Math Performance, Grade Level, and Spacing Conditions

There is a common view that students' computational errors are an indication of their conceptual misunderstandings about mathematics. Consistent with this idea, students' prior performance in ASSISTments significantly predicted their mastery speed, suggesting that, on average, students with higher prior performance made fewer errors when solving order of operations problems. However, even when controlling for prior performance, incongruent spacing still affected student performance on the problem set. Additionally, while previous findings have largely focused on college-aged participants in laboratory settings (e.g., Gómez, Bossi, & Dartnell, 2014; Jiang et al., 2013; Landy & Goldstone, 2007a, 2007b, 2010), the current findings reveal that the effects of physical spacing also occur in younger students who are in the process of learning the rules of operation precedence and applying that knowledge in authentic homework settings. Regardless of age or prior performance in the tutoring system, viewing expressions with spacing that is incongruent with the order of precedence seems to be more difficult for students ranging from the 5th to 12th grade.

Taken together, these findings reinforce previous evidence that subtle changes in physical spacing can impact students' performance on computing order of operations problems regardless of the student's age or knowledge level. It seems that perceptual grouping, through spacing, may be acting as an irrelevant but substantial lure that is hard for students to ignore. More broadly, the differences in performance across conditions supports the notion that people use space as a perceptual cue when interpreting and acting on mathematical symbols. As such, these results provide further evidence that visual and perceptual processes can drive reasoning about mathematics computations (Landy & Goldstone, 2007a, 2007b) and highlight the conflicts between relevant (rule-ordered) and irrelevant (spacing) features in the presentation of mathematics.

Implications for Teaching and Learning

The current study suggests that minor and relatively meaningless changes to the visual presentation of mathematical notation have implications for how students interpret and use symbolic notation to perform computations. Although perceptual cues influence mathematical reasoning, few instructional approaches or interventions make use of the power of perception. Future learning interventions for algebra learners could include purposeful manipulations to the presentation of mathematical expressions and equations which could affect students' abilities to learn and apply arbitrary mathematical rules such as the order of precedence. More broadly, the prevalence of a spacing effect on mathematics performance across upper elementary through high school age groups poses interesting questions and may have theoretical implications about when perceptual cues begin to drive cognitive processes in learning and development. Future work could investigate how early spacing effects emerge in young learners and how spatial manipulations may drive students' cognitive processes and actions at younger ages.

Limitations

There are several limitations to the dataset as well as the methods used in these analyses. For instance, the Skill Builder structure (where students must correctly answer three problems in a row) is not a common approach used in classroom instruction and does not easily lend itself to a pretest/posttest design. The Skill

Builder structure also allows students to stop working after they achieve mastery status in an assignment by correctly answering three problems in a row. As a result, participants only answered an average of six problems in the assignment. Additionally, since the final dataset excluded students who answered the first three problems correctly, this sample does not take into account the behavior of the highest-performing mathematics students on this particular problem set.

Another limitation of the study is that there was limited demographic data available on the students. Since ASSISTments problem sets were assigned by teachers around the country who use the platform for homework, we are unable to collect specific data about children, teachers, or the classroom context. While this is certainly a limitation, the fact that ASSISTments Skill Builders are used and assigned widely by teachers allowed for more ecological validity and a much larger sample size than would have been collected if this study was conducted by researchers with recruited teachers in local classrooms. Additionally, effects of spacing were found even while controlling for prior performance and grade level.

Future Directions

While the current project focused solely on assignment mastery, subtle spacing manipulations have been found to influence student problem-solving behavior at the action level (Jiang et al., 2013). Based on work such as that of Jiang and colleagues (2013, 2014), examining other aspects of behavior while students are problem-solving may be fruitful lines of research. Jiang and colleagues (2013) manipulated perceptual grouping in a similar study using the minus sign ("-") as the operand of interest then analyzed rates of "target errors" that would be the result of relying on perceptual grouping rather than arithmetic precedence (i.e. subtracting 14 from "20 - 4 + 10"). Jiang et al. (2013) found that physical spacing manipulations did, in fact, lead to higher rates of target errors. Investigating the kinds of mistakes that are made on incorrect problems could provide new insights about the role of spacing in expressions and how perceptual groupings drive student actions while problem-solving.

One difference in our study from prior work is that we included a mixed spacing condition. While findings suggest that the mixed spacing condition was more difficult for students compared to students in the congruent spacing condition, interestingly, there were no differences in mastery speed between the mixed condition and the neutral condition, or the mixed and incongruent spacing conditions. Given these mixed findings, future experimental work should further examine the effects of mixed spacing and test plausible mechanisms for the impact of perceptual spacing.

It is also important to note that, across conditions, students were provided with an optional hint on each problem to remind them of the rules of precedence. Roughly a third of participants viewed the hint at least once while working on the assignment. That said, little to no work has studied how conceptual knowledge reminders about the order of precedence may mitigate the effect of physical spacing on mathematics performance. To develop a richer understanding of how perceptual grouping may affect student behavior, our team is currently exploring patterns of student behavior within problems, such as response times, help-seeking behavior, and error types, to see how effects of spacing manipulations extend to a broader population of students. More thorough error analyses at the action level, within problems, could provide insight about whether students who are presented with incongruent spacing tend to make predictable errors based on how the symbols were

visually grouped. We also plan to explore students' actions after viewing the available hint to examine whether order of precedence errors continue to occur after a visual reminder of the order of precedence.

Conclusion

This work demonstrates that irrelevant, but salient visual information in notation, such as spacing, can influence students' reasoning of mathematics. Specifically, perceptual cues, even those that are mathematically misleading such as incongruent spacing, are difficult to ignore. This study extends prior work in the following ways. First, to our knowledge, this is one of the first randomized controlled trials conducted on physical spacing in the context of an online learning platform with school aged children, showing that perceptual grouping occurs in authentic learning environments in addition to laboratory settings. Second, these results reveal that incongruent spacing between terms does impact 5th to 12th grade students' performance on mathematics assignments, resulting in slower mastery speeds, even when a reminder about the order of precedence is available as a hint on each problem. Conversely, terms in mathematical expressions that are neutrally spaced or grouped together to be congruent with the order of precedence increase assignment mastery speeds. These findings further support the notion that subtle perceptual cues, such as physical spacing, do not bear any practical implications in mathematics yet have effects on mathematical cognition and performance for students in upper elementary through high school. Learners and experts alike utilize perceptual strategies when reasoning about mathematics.

Competing Interests

We have no competing interests to disclose.

Data Availability

For this study, a dataset is freely available (Harrison, Smith, Hulse, & Ottmar, 2020b).

Supplementary Materials

The de-identified data used in this analysis is available on the Open Science Framework. Additionally, the original data report from ASSISTments is available for further reference (for access, see Index of Supplementary Materials below).

PsychOpen GOLD

Index of Supplementary Materials

Harrison, A., Smith, H., Hulse, T., & Ottmar, E. R. (2020a). *Supplementary materials to "Spacing out! Manipulating spatial features in mathematical expressions affects performance"* [Original data report]. ASSISTments. https://tiny.cc/spacingdata

Harrison, A., Smith, H., Hulse, T., & Ottmar, E. R. (2020b). *Supplementary materials to "Spacing out! Manipulating spatial features in mathematical expressions affects performance"* [Research data]. OSF. https://doi.org/10.17605/OSF.IO/BAEUJ

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY, USA: Oxford University Press.

Botelho, A., Wan, H., & Heffernan, N. (2015, March). The prediction of student first response using prerequisite skills. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale* (pp. 39-45). New York, NY, USA: ACM.

Braithwaite, D. W., Goldstone, R. L., van der Maas, H. L. J., & Landy, D. H. (2016). Non-formal mechanisms in mathematical cognitive development: The case of arithmetic. *Cognition, 149*, 40-55. https://doi.org/10.1016/j.cognition.2016.01.004

Gibson, E. J. (1969). *Principles of perceptual learning and development*. East Norwalk, CT, USA: Appleton-Century-Crofts.

Goldstone, R. L., Marghetis, T., Weitnauer, E., Ottmar, E. R., & Landy, D. (2017). Adapting perception, action, and technology for mathematical reasoning. *Current Directions in Psychological Science, 26*(5), 434-441. https://doi.org/10.1177/0963721417704888

Gómez, D. M., Benavides-Varela, S., Picciano, C., Semenza, C., & Dartnell, P. (2014). Symbol spacing and its influence in school children's calculation of mixed-operation arithmetic expressions. In S. Oesterle, C. Nicol, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 6, pp. 310). Vancouver, Canada: PME.

Gómez, D. M., Bossi, B., & Dartnell, P. (2014). Spacing patterns affect students' application of arithmetic precedence. In S. Oesterle, C. Nicol, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 6, pp. 310). Vancouver, Canada: PME.

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*, 470-497. https://doi.org/10.1007/s40593-014-0024-x

Jiang, M. J., Cooper, J. L., & Alibali, M. W. (2013). Visually grouping operands: Perceptual factors influence arithmetic performance. In M. Martinez & A. Castro Superfine (Eds.), *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 253-259). Chicago, IL, USA: University of Illinois.

Jiang, M. J., Cooper, J. L., & Alibali, M. W. (2014). Spatial factors influence arithmetic performance: The case of the minus sign. *Quarterly Journal of Experimental Psychology, 67*(8), 1626-1642. https://doi.org/10.1080/17470218.2014.898669

Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics Education, 20*, 274-287. https://doi.org/10.2307/749516

Landy, D., & Goldstone, R. L. (2007a). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 720-733. https://doi.org/10.1037/0278-7393.33.4.720

Landy, D., & Goldstone, R. L. (2007b). Formal notations are diagrams: Evidence from a production task. *Memory & Cognition, 35*(8), 2033-2040. https://doi.org/10.3758/BF03192935

Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly Journal of Experimental Psychology, 63*(10), 1953-1968. https://doi.org/10.1080/17470211003787619

Landy, D. H., Jones, M. N., & Goldstone, R. L. (2008). How the appearance of an operator affects its formal precedence. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 2109–2114). Washington, DC, USA: Cognitive Science Society.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist, 35*(2), 125-141. https://doi.org/10.1207/S15326985EP3502_6

Marghetis, T., Landy, D., & Goldstone, R. L. (2016). Mastering algebra retrains the visual system to perceive hierarchical structure in equations. *Cognitive Research: Principles and Implications, 1*(1), Article 25. https://doi.org/10.1186/s41235-016-0020-9

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466. https://doi.org/10.1207/s15516709cog2803_7

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899. https://doi.org/10.1111/j.1467-8624.2005.00884.x

Niehaus, E., Campbell, C. M., & Inkelas, K. K. (2014). HLM behind the curtain: Unveiling decisions behind the use and interpretation of HLM in higher education research. *Research in Higher Education, 55*, 101-122. https://doi.org/10.1007/s11162-013-9306-7

Ostrow, K. S., Selent, D., Wang, Y., VanInwegen, E. G., Heffernan, N. T., & Williams, J. J. (2016). The assessment of learning infrastructure (ALI): The theory, practice and scalability of automated assessment. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 279-288). New York, NY, USA: ACM.

Rivera, J., & Garrigan, P. (2016). Persistent perceptual grouping effects in the evaluation of simple arithmetic expressions. *Memory & Cognition, 44*(5), 750-761. https://doi.org/10.3758/s13421-016-0593-z

Rumelhart, D., Smolensky, P., McClelland, D., & Hinton, G. (1986). Schemata and sequential thought processes in PDP models. In *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 7–57). Cambridge, MA, USA: MIT Press.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*, 1172-1217. https://doi.org/10.1037/a0029333

PsychOpen GOLD

Walkington, C., Clinton, V., & Sparks, A. (2019). The effect of language modification of mathematics story problems on problem-solving in online homework. *Instructional Science, 47*, 499-529. https://doi.org/10.1007/s11251-019-09481-6

Wertheimer, M. (1950). Gestalt theory. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 1–11). New York, NY, USA: Humanities Press.