Effects of Problem Order on Accuracy, Preference, and Choice of Multiplication Assignments

Kathryn E. Jaspers, Christopher H. Skinner, Robert L. Williams & Lee B. Saecker

College students (*N* = 151) completed three in-class mathematics assignments that varied with respect to the order of short, medium, and long multiplication problems, including a short-to-long order, a long-to-short order, and a random order. Problem order within assignments did not significantly affect work accuracy, completion time, or assignment preference. However, participants rated (1) the assignment with random order as less difficult than the other two assignments and (2) the assignment with long-to-short order as requiring less time than the other two assignments. Participants were likely to choose a related homework assignment viewed as less effortful, less difficult, or less time consuming. Several concepts (e.g., matching law, behavioral momentum, and Premack reinforcement principle) were used in interpreting the results.
Keywords: problem order, mathematics, homework, matching law, behavioral momentum, Premack reinforcement principle.

Academic assignments and tests often involve a series of tasks or problems that can be ordered in a variety of ways. A typical pattern is to begin the assignment with the less demanding tasks (briefer, easier) and proceed to the more demanding. A less common order is to begin with the more demanding tasks and proceed to the less demanding. A third order is to randomize task order with respect to difficulty. Because order of tasks within an assignment could affect performance on the assignment, time to complete the assignment, and favorability rating of the assignment, one would expect a plethora of research dealing with the effect of task order on student performance and preference of task order. However, an examination of research on assignment performance and assignment preference reveals few studies that deal specifically with the effects of task order within assignments.

Research in this area has focused mainly on exams rather than assignments. Studies of multiple-choice exams have demonstrated mixed effects of item order on performance accuracy. Balch (1989) found that sequential exams (in which the test material was presented in the same order that material was presented in texts or lectures) produced higher scores than randomized or chapter-sequenced exams. In contrast, Neely, Springston, and McCann (1994) reported no differences in accuracy between sequential and randomized questions on undergraduate psychology exams. Similarly, Perlini, Lind, and Zumbo (1998) found no influence of easy-to-hard, hard-to-easy, or random item order on psychology exam performance.

Studies of difficulty ratings of exams with different task orders have also produced mixed results. Laffitte (1984) found no differences in perceived difficulty of exams organized as easy-to-hard by chapter, easy-to-hard across chapters, randomly within chapters, and randomly across chapters. In contrast, Olson and Barickowski (1976) reported that hard-to-easy tests were viewed as easier than easy-to-hard, with no differences in accuracy of answers. Whether these research findings on item order with psychology exams can be generalized to assignments that involve skill building in a variety of academic domains remains an issue for further investigation.

Research regarding the effects of choosing assignment sequence has been done mainly with children and adults with special needs (e.g., autism, mental retardation) and has typically dealt with the scheduling of assignments rather that the order of tasks within assignments. For example, Kern, Mantegna, Vorndran, Bailin, and Hilt (2001) found that students having intellectual and behavioral problems improved both in task engagement and appropriate conduct when allowed to schedule the sequence of required tasks. Similarly, Seybert, Dunlap, and Ferro (1996) reported that giving high school students with severe mental retardation the opportunity to schedule their tasks increased task engagement

and task productivity. Watanabe and Sturmey (2003) found that even adults with autism spent more time on task when allowed to schedule their work activities. The study closest in choice procedures to the current study found that giving children with autism an opportunity to select the order of homework activities and the order of tasks within activities increased homework completion, correctness of homework, and positive affective ratings (Moes, 1998).

Irrespective of the choice of assignment or task sequence, one would expect completion of an assignment to function as a reinforcing event for students. Skinner's (2002) meta-analysis of interspersal research posits that a completed discrete task within an assignment also functions as a reinforcing event. If assignments are equated for number and difficulty of tasks, one would assume these assignments to offer equivalent potential for reinforcement derived from completion of discrete tasks. According to this notion, task order within assignments should not affect the cumulative experience of reinforcement within assignments containing the same number of tasks and the same range of task difficulty.

A conception that extends Skinner's (2002) hypothesis of reinforcement from discrete task completion is Herrnstein's (1961) matching law, which stipulates that when students are given the choice of two behaviors, all else being equal, they are likely to choose the behavior associated with stronger reinforcement (i.e., more immediate, higher quality, or thicker schedule). However, given the same number and type of tasks across assignments, the major reinforcement parameters would appear to be held constant across assignments. Nonetheless, it is possible that task sequence within assignments could affect students' perception of assignment difficulty. That being a possibility, one would expect students subsequently to choose the task order they initially perceived as requiring the least effort. Thus, researchers first must establish which task order students perceive as least and most difficult before applying the matching law in predicting which task order students will prefer for future assignments.

A perspective that may be more illuminating than either the discrete task completion hypothesis or the matching law in predicting the effect of task order on assignment performance and preference is the notion of behavioral momentum. Behavioral momentum has been described as a behavioral manifestation of Newton's second law of motion (Nevin, 1992; Nevin, Mandell, & Atak, 1983; Plaud & Gaither, 1996). According to Newton's 2nd law, acceleration or momentum is the result of an object's mass and the force pushing the object. Some behavioral researchers embracing the notion of behavioral momentum have proposed that Newton's 2nd law accounts for task performance in much the same way it explains physical phenomena. Contingencies (force) cause an increase in response rate (velocity), which is necessary to overcome resistance (mass) to change (Nevin, 1992; Nevin, 1996; Nevin et al., 1983). Thus, performing a series of high-probability actions provides momentum to overcome resistance inherent in performing lower probability actions, which typically are more demanding tasks (Nevin, 1996). Consequently, a high rate of reinforcement in the early stages of task performance may sustain one's task completion when the attainment of reinforcement becomes more arduous (Lee et al., 2006).

The Premack reinforcement principle (Premack, 1959) provides yet another perspective on task completion within assignments. According to this principle, completion of a lower probability action can be reinforced by subsequent access to a higher probability action. Thus, if tasks within an assignment are ordered from most to least demanding, the completion of each task would be followed by a less demanding task. Presumably, the impetus to complete each task would be bolstered by the prospect of a subsequent less-demanding task. Another potential advantage of this sequence would be that one would address the most difficult tasks when one's cognitive energy level is highest (i.e., at the beginning of the assignment).

Framework for the Current Study

All task sequences have potential advantages and disadvantages according to the behavioral momentum notion and the Premack reinforcement principle. The sequence from short to long should

build momentum as one progresses through an assignment but may diminish student choice of that task sequence for a future assignment, given that the initial assignment finishes with the most difficult tasks. Alternatively, the Premack reinforcement principle suggests that the long-to-short sequence may start slowly as one tackles the most difficult problems at the outset of an assignment, but one's effort on the difficult problems will be reinforced by progression to less difficult tasks. Also, given that the task sequence finishes with the least difficult tasks, the long-to-short sequence may become the student's preferred order for a future assignment.

In contrast to the implications of the behavioral momentum concept and the Premack reinforcement principle, Skinner's (2002) conception of discrete task reinforcement appears to minimize the influence of task order per se on progression through an assignment or preference for a particular task order in future assignments. Likewise, the matching law would seem to suggest that problem sequence has minimal effect on assignment preferences and choices of subsequent assignments when the same number of long, medium, and short problems is included in each assignment. Presumably, within the framework of the matching law, perceptions and choice are based on the average relative rates of reinforcement throughout the assignment. Nonetheless, if various task orders are perceived by students as requiring differential effort, the matching law would imply that students will subsequently choose the task order perceived as requiring the least effort.

Within the combined conceptual framework of the behavioral momentum concept, the Premack reinforcement principle, the discrete task reinforcement hypothesis, and the matching law, the current study compared the effects of three sequences of math problems within assignments (i.e., short to long, long to short, and random) on several dependent measures: accuracy of problem solutions, time required to complete each assignment, perceived difficulty of each assignment, estimated time required to complete each assignment, assignment preference, and choice of a related homework assignment. It was assumed that the task order might have greater affects on students' perceptions of assignments than actual performance and time spent on assignments.

Method

Participants

Participants for this study were students in an undergraduate human development course in the Fall 2006 semester at a large public university in the Southeastern US. Students received 5 points of credit for their participation, which was less than 1% of their final grade. Three class sections, 50 to 51 students per section, were invited to participate. Although 151 students elected to participate, six did not follow directions (e.g., skipped problems) and three did not complete each assignment in the allotted time. Thus, data were analyzed for 142 students. Nine students who failed to turn in their homework or did not correctly complete the homework assignment were included for their in-class assignments but not for their homework accuracy. Approximately 76% of the participants were female. For academic classification, 4% reported they were freshman, 40% sophomores, 26% juniors, 16% seniors, and 3% graduate students. About 2% of the sample reported "other" as their academic classification and 9% did not report their academic level.

Materials

Students were asked to work mathematics problems from three types of assignments varying only in the order of problems within assignments. Each assignment was presented on one side of 8.5 x 11 inch white paper. Each began with a blank in which the students wrote their start time, followed by nine multiplication problems, then a blank space for their end time, and finally a direction for the next action (e.g., "go to the next page," "stop").

Assignment types. Researchers constructed three mathematics sheets differing in the order in

which nine multiplication problems were to be completed. In the short-to-long condition, the participants began with three 2 X 1 (e.g., 65 x 4) problems, followed by three 2 X 2 (e.g., 69 x 57) and three 3 X 3 (e.g., 864 x 987). The second condition began with longer problems and progressed to shorter problems. In the third condition, six versions of a random sequence were developed, with each sequence randomly distributed to the same number of students. All random orders were first generated using a random-numbers sequence, but then adjusted to make sure they started and ended with each length (short, medium, long) represented an equal number of times. Plus, no more than two problems of the same length were permitted to occur in succession. The intent of developing different versions of the randomized assignment was to ensure that no specific aspect of the randomized assignment dominated the students' perception of the random sequence. For example, if all randomized assignments had ended with a brief problem, students might have developed an unduly favorable view of the randomized assignment.

All multiplication problems had similar characteristics: all used digits over 4 in order to ensure that every computation involved carrying, but 5s were omitted because students often learn to multiply by 5s quickly and automatically. In the 2 X 1 and 2 X 2 problems, no digits were repeated. In the 3 x 3 problems, one digit was repeated because there were only five unique digits (4, 6, 7, 8, and 9) and 6 digits per problem. For each assignment, the problems were not numbered and were presented in an unbalanced format (i.e., unequal spacing and unequal number of problems per row). Problems were matched across assignments with respect to difficulty level and specific operations required. For example, if Assignment J had 96 x 7, then the next assignment would have 97 x 6 and the third assignment, 76 x 9. This ensured that each problem would involve the same computations across assignments, thus controlling for problem difficulty.

Experimental packets.

Packets were assembled by counterbalancing the order of the three assignments. Each assignment sequence was randomly distributed to the same number of students. Packets consisted of a cover page with directions and an informed consent, followed by an assignment, then a demographics page to allow time for dissipation of the cognitive residual of completing the first assignment. The demographics page included spaces for students to indicate their age, sex, and race. A second assignment followed the demographics page, and another demographic page following the second assignment. The second demographic page requested more information from the students (such as year in school, the most recent math class taken, the semester and year of most recent math class, and their approximate GPA). The third and final assignment followed, and then the packet ended with a page of questions concerning the math assignments the students had just completed. These questions asked which assignment was most difficult, least difficult, took the most effort, took the least effort, required the most time, required the least time, and finally, which assignment students wanted to do for homework. The homework assignments mirrored the in-class assignment in length of problems, number of problems, and construction, although none of the problems were the same as those in the in-class assignment. Only one version of the randomized assignment was developed as a homework assignment, rather than six versions as were constructed for the in-class assignment.

Procedure

All experimental procedures took place within the students' regular classroom during scheduled class time on the second day of class. Upon submitting signed informed consent forms, students completed their packet including the three assignments, the two demographic questionnaires which separated the assignments, and the sheet for evaluating the assignments. Before beginning the assignments, students were instructed to complete all problems in order—working from left to right and top to bottom. This sequence was demonstrated to the students using an overhead of a sample assignment. Additionally, students were told to answer all questions on the demographic and evaluation pages.

As students worked on the packet, four research assistants moved around the room to monitor students' adherence to task instructions. If students worked problems out of sequence (e.g., skipped problems), their data were excluded from analysis. A large screen was used to depict the time in seconds. Students used this screen to record their starting and stopping time for each assignment. The cover page included directions for writing the starting and stopping time. This procedure also was demonstrated to the students prior to beginning the assignments. After 35 minutes, all students' materials were collected, even from students who had not finished their packets.

Dependent Variables and Data Analysis Procedures

Several dependent variables were analyzed including number of problems correct; time to complete assignments; ranking of assignments according to perceived time, difficulty, and effort; and choice of a related homework assignment. We excluded students who did not complete all three assignments; thus, accuracy scores could range from 0-9 for each assignment. Time spent on each assignment was calculated by subtracting the finishing number of seconds noted by the student from the starting number of seconds indicated by the student.

After finishing the three assignments, participants completed a seven-question evaluation sheet, on which they rated the assignments as to which assignment was perceived as requiring the most effort and least effort, was most difficult and least difficult, and took the most time and least amount of time. Finally, participants chose one homework assignment from three assignments reflecting the different task orders.

Analysis and interscorer agreement.

One of two researchers calculated the accuracy of and time to complete each assignment. These determinations were made from the packets submitted by the students at the completion of the in-class research activity. In order to establish interscorer agreement, each researcher scored 20% of the other's assignments. Interscorer agreement was 99.6%.

Procedural Integrity.

A treatment-integrity sheet was developed to ensure that all directions were followed. Included on the integrity sheet was the following steps: set up the clock; have the example overhead ready; designate a graduate student to pick up packets; read directions aloud from cover sheet; ask if students have any questions; have students sign informed consent form; instruct students to turn the page, write beginning time, and begin assignments; and start timer. An independent research assistant observed sessions and noted each behavior on the integrity sheet. Procedural integrity proved to be 100%.

Results

The researchers used repeated measures, chi-square, and cross-tab analyses to determine the relationship between problem order and a variety of variables, including accuracy of problem solutions, time to complete each assignment, differential perceptions of assignments, and homework choice. Chi-square rather than correlational analyses were used because of the categorical nature of the perceptual variables. An alpha level of .05 was used to determine significance for all statistical tests.

Assignment Order

Accuracy and completion time. The initial analysis determined whether assignment order per se affected accuracy, time to complete assignments, and perceptions of assignments. Participants' mean number of problems correct on the first, second, and third assignments were 6.40, 6.56, and 6.52 problems correct, respectively. A repeated measures analysis showed that participants' accuracy did not differ regardless of the order in which the assignments were completed, $F(2,140) = .57$, $p > .05$. Participants spent an average of 6.67 minutes completing the first assignment, 6.00 minutes to complete

the second assignment, and 5.63 minutes to complete the third assignment. These means were significantly different, $F(2, 140) = 14.73$, $p < .01$ (see Table 1). More specifically, time spent on the first assignment was significantly more than on the second assignment, $t(132) = 3.33$, $p < .01$, and third assignment, $t(132) = 5.04$, $p < .001$, and time spent on the second assignment was significantly greater than time spent on the third assignment, $t(131) = 3.22$, $p < .01$.

Table 1
Assignment Order: Completion Time and Assignment Accuracy

Assignment Order	*n*	*M*	*SD*
	Completion time[a]		
1st assignment	136	6.67	2.67
2nd assignment	136	6.00	2.04
3rd assignment	136	5.63	1.85
	Assignment accuracy[b ns]		
1st assignment	142	6.40	1.90
2nd assignment	142	6.56	1.90
3rd assignment	142	6.52	1.98

[a]in minutes, these means were significantly different at $p < .01$.
[b]number of problems correct (out of 9 possible) for in-class assignments. [ns]not significant.

Perception of assignments.

Assignment order was related to perceptions of assignment difficulty, time to complete assignments, and effort required to complete assignments. Participants perceived their first assignment, irrespective of problem sequence, as the being the most difficult, $\chi^2(1, n = 140) = 38.62$, $p < .01$, taking the most time, $\chi^2(1, n = 140) = 35.85$, $p < .01$, and requiring the most effort, $\chi^2(1, n = 141) = 26.31$, $p < .05$. Although the first assignment was perceived as more difficult, more effortful, and taking more time, the converse was not true: the last assignment was *not* perceived as least difficult, $\chi^2(1, n = 140) = 14.41$, $p > .05$., least effortful, $\chi^2(1, n = 141) = 14.10$, $p > .05$, or requiring less time, $\chi^2(1, n = 142) = 13.81$, $p > .05$. In addition, assignment order did not affect homework choice, $F(1, 5) = 1.52$, $p > .05$.

Assignment Type

Accuracy and completion time. Participants' mean number of problems correct on the short-to-long, long-to-short, and randomly sequenced assignments in class were 6.35, 6.65, and 6.49 problems correct (out of 9), respectively. Repeated measures analysis showed these means not to differ significantly, $F(2, 140) = 2.23$, $p > .05$. Participants spent an average of 5.86 minutes to complete the short-to-long assignment, 6.04 minutes to complete the long-to-short assignment, and 6.24 minutes to complete the randomly sequenced assignment, but again these differences were not significant, $F(2, 128) = 1.53$, $p > .05$ (see Table 2).

Table 2
Assignment Type: Completion Time, Assignment Accuracy, and Homework Accuracy

Assignment Type	*n*	*M*	*SD*
	Completion time[a ns]		
Short-to-long	134	5.85	2.04
Long-to-short	138	6.14	2.30
Random	136	6.32	2.39
	Assignment accuracy[b ns]		
Short-to-long	142	6.35	1.98
Long-to-short	142	6.65	1.86
Random	142	6.49	2.30
	Homework accuracy[c ns]		
Short-to-long	46	8.20	1.02
Long-to-short	46	7.89	1.29
Random	40	7.75	1.74

[a]in minutes.
[b]number of problems correct (out of 9 possible) for in-class assignments.
[c]number of problems correct (out of 9) for chosen homework assignment.
[ns]not significant.

Perception of assignments.

Participants perceived the short-to-long assignment as more difficult than the randomly sequenced assignment, $\chi^2(1, n = 92) = 6.26$, $p = .01$. The randomly sequenced assignment was rated as less difficult than short-to-long or long-to-short assignments, $\chi^2(1, n = 102) = 4.75$, $p = .03$ and $\chi^2(1, n = 95) = 5.76$, $p = .02$, respectively. The randomly sequenced assignment also was rated as less effortful than the long-to-short assignment, $\chi^2(1, n = 97) = 6.44$, $p = .01$. Participants judged the short-to-long and randomly sequenced assignments as taking more time than long-to-short assignment, $\chi^2(1, N = 89) = 7.02$, $p = .01$ and $\chi^2(1, N = 83) = 4.35$, $p = .04$, respectively (see Table 3). Additional general linear model multivariate analyses showed that neither assignment accuracy nor completion time was related to the perception of effort or difficulty.

Table 3
Choice and Perceived Difficulty, Time, and Effort

Assignment Type	*Short-to-long* Freq.	(%)	*Long-to-short* Freq.	(%)	*Random* Freq.	(%)
Most difficult[a]	58	(40.8)	48	(33.8)	34	(23.9)
Least difficult[b]	40	(28.3)	38	(26.8)	62	(43.7)
Most effort[ns]	56	(39.4)	46	(32.4)	39	(27.5)
Least effort[c]	44	(31.0)	36	(25.4)	61	(43.0)
Most time[d]	57	(40.1)	32	(22.5)	51	(35.9)
Least time[ns]	47	(33.1)	40	(28.2)	55	(38.7)
Homework choice[ns]	48	(33.8)	50	(35.2)	44	(31.0)

[a]Short-to-long rated more difficult than random (p = .01).
[b]Random rated as less difficult than short-to-long (p = .03) and long-to-short (p = .02).
[c]Random rated as less effortful than long-to-short (p = .01).
[d]Short-to-long (p = .01) and random (.04) rated as taking more time than long-to-short [ns]not significant

Choice of homework.

Participants did not choose any type of homework assignment more frequently than any other type, $\chi^2(2, n = 142) = .39$, $p > .05$ (see Table 3), yet participants were more likely to choose an assignment they perceived as less difficult, $\chi^2(1, N = 140) = 19.76$, $p < .01$, less effortful $\chi^2(1, N = 140) = 23.25$, $p < .01$, or less time consuming, $\chi^2(1, N = 140) = 4.57$, $p < .01$, for the homework assignment. As was the case for the perceptual ratings of assignments, neither accuracy nor completion time was related to choice of homework assignment. Although ancillary to the purposes of the study, a gender difference was found in choice of homework assignments. Male participants were more likely to choose an assignment they perceived as more effortful and difficult, whereas female participants were more likely to choose an assignment they perceived as requiring less effort and as being less difficult, $\chi^2(1, N = 87) = 13.68$, $p < .01$.

One might expect that task order would have its greatest effect on homework choice when a particular task order occurred in the last homework assignment for a particular student. For example, the long-to-short sequence might contribute more to a student's choice of a long-to-short homework assignment if the long-to-short sequence occurred as the last assignment rather than the first or middle assignment. Any momentum effect from ending the assignment on shorter problems would probably be lost if this order was embedded in the first or second assignment. Chi-squared analyses indicated there were no differences between last assignment type and homework choice, $\chi^2(2, N = 142) = .502$, $p = .78$. This finding suggests that not even the last assignment completed before choosing a homework assignment affected students' choice of a homework assignment.

Discussion

The study compared three assignments composed of nine virtually identical problems in different orders (long-to-short, short-to-long, and random). Thus, the total amount of discrete-task reinforcement and effort required to complete assignments should have been equivalent across assignments. Given this assumption, both the discrete task reinforcement hypothesis and the matching law (Herrnstein, 1961)

would suggest no difference in any of the dependent variables across assignment type. This proved to be the case for performance variables (accuracy and completion time) but not for perception of assignments. The assignment with random order of tasks was generally perceived as less difficult and less effortful than the other assignment types. Thus, one might have expected the random sequence to be chosen most frequently for the homework assignment. Although this did not prove to be the case, the analysis did show that in-class assignments generally perceived as less difficult, less effortful, or less time consuming were chosen most frequently for homework. This latter finding would seem to be consistent with the matching law.

Neither the behavioral momentum nor the Premack reinforcement notion was strongly supported by the findings. The three task orders did not differ in their effects on completion time or accuracy of work. However, perception of assignments showed that the short-to-long sequence, which was most consistent with the behavioral momentum notion, tended to be rated less favorably (most difficult, most effortful, and most time consuming) than the other task orders. In predicting homework choice, one would expect the long-to-short sequence to be most consistent with behavioral momentum, inasmuch as the in-class assignment finished with several shorter problems. Presumably, a primacy perspective might cause the ending problems to be figural in student choices of a homework assignment. However, no significant differences or apparent trends emerged in homework choices between long-to-short and the other assignment types.

The long-to-short sequence (consistent with the Premack reinforcement principle) fared little better than the short-to-long sequence in differentiating the findings for task sequence. In fact, the lowest percentage of students regarded the long-to-short sequence as least difficult, least effortful, and least time consuming compared to other assignment types. Paradoxically, the lowest percentage of students also rated the long-to-short sequence as most difficult, most effortful, and most time consuming, though not all of these differences were statistically significant. Nonetheless, the raw-score differences between long-to-short and short-to-long sequences were greater for the most difficult, most effortful, and most time consuming ratings than for the least difficult, least effortful, and least time consuming ratings. Hence, there may be some suggestion that students generally find the long-to-short sequence more palatable than the short-to-long sequence.

The apparent winner in favorability ratings for task order within assignments appears to be the random sequence. Although not all percentage differences between this task order and other task orders proved statistically significant, a smaller percentage of students rated the random sequence as most difficult, most effortful, and most time consuming compared to student ratings of other task orders. Conversely, a higher percentage of students rated the random order as least difficult, least effortful, and least time consuming compared to other task orders. However, this favorability advantage did not transfer to choice of homework assignments, with the random sequence chosen by a similar percentage of students as the other task sequences.

What might account for the finding that students rated the random order more acceptable than assignments that have directional problem order proceeding from short to long or long to short? A potential problem with the short-to-long sequence is that students can see from the outset of the assignment that problems will get longer, and a potential problem with the long-to-short sequence is that the assignment begins at its most difficult level. Plus, more frequent changes in difficulty level may provide students more distributed reinforcement from working the shorter problems than grouping these problems at the beginning or end of a task sequence.

Perceived differences in in-class assignments, even though they did not strongly favor any one type of task order, led to discernible patterns in homework choices. Participants were more likely to choose a homework assignment they viewed as less effortful, less difficult, or requiring less time.

Apparently, there is enough individuality in task-sequence preferences to make it difficult to predict what task sequence would be preferred by students in general. However, once teachers determine each individual's preferred task sequence, that information has increased predictive potential. However, the data arrangement in the current study did not permit a determination of whether students performed better or more efficiently on their most preferred homework assignment than on their least preferred, inasmuch as all students worked on the assignment they chose and did their homework out of class under unmonitored conditions.

The findings of this study do not permit claims of substantial superiority of any one task sequence over the other task sequences, contrary to both the behavioral momentum and Premack reinforcement notions. Nonetheless, some methodological arrangements in the current study may limit both the internal and external validity of the findings. First, the assignments were very short with only three levels of problem length within assignments. Longer assignments with greater variation in problem length might provide a stronger test of the task-sequence effect. Only three levels of problem length were used in the current study because many teachers appear to group tasks by levels rather than differentiating length or difficulty across all tasks. Nonetheless, having only three levels may have weakened the test of both behavioral momentum and the Premack reinforcement principle.

A second methodological limitation was the use of a 3-choice design, which may have increased the difficulty of choosing among alternative assignments. Choice research (e.g., Skinner, 2002) typically has required a choice between two behaviors (e.g., Myers & Myers, 1977). It may be that choosing between three very similar assignments (same number of problems and same range of difficulty) blurred the distinctions between options and made student choices less valid and reliable. Further research using a 2-choice comparison between traditional short-to-long assignments and less common long-to-short assignments might allow for a clearer distinction between these sequences.

A third possible limitation was that the 2 X 1 ("brief") problems may not have been brief enough to adequately test the behavioral momentum notion. The theory of behavioral momentum requires a series of high probability behaviors before a lower probability behavior is required, with the high probability behaviors tending to be virtually automatic responses. Admittedly, the brief problems in the various assignments should have been quite easy for college students, but they did require some level of math computation. Consequently, the brief problems may not have produced the automatic responses that presumably occur early in a momentum sequence.

Additionally, most of the behavioral momentum research has focused on increasing *compliance* with a low-probability task, oftentimes by first introducing unrelated high-probability responses (see Lee, 2005, for a meta-analysis of behavioral momentum research). In our study, we operated under the assumption that one manifestation of behavioral momentum would be students' choice of a long-to-short homework assignment after completing an in-class assignment ending with three high-probability problems. However, any delay between the high- and low-probability behaviors likely weakens behavioral momentum. Lee (2005) concluded that interventions that rely on behavioral momentum are most effective when the time between the final high-probability response and the low-probability request is less than 10 sec.

Another factor that may have undermined the possibility of a task-sequence effect within assignments was the apparent strength of an assignment-sequence effect. Because the first assignment was reported as more difficult, effortful, and time-consuming than the other assignments, participants may have had difficulty in distinguishing comparable differences among task orders. Thus, assignment order may have overpowered the perception of task order within assignments. The students were asked to engage in a behavior (long multiplication with carrying) that perhaps many had not done in years without

the aid of a calculator. Although students were not less accurate on the first assignment, that assignment may have required more cognitive investment than the other assignments.

Gender also appears to have affected how students perceived the difficulty of math assignments. In the current study, male participants were more likely to choose an assignment they perceived as more difficult and effortful, whereas female participants were more likely to choose an assignment that they perceived as less difficult and effortful. Although males were more likely than females to choose an assignment perceived as difficult and effortful, they were not faster or more accurate on the in-class assignments. Previous research has shown differences between males and females in mathematics regarding self-efficacy (Junge & Dretzke, 1995) or confidence (Lloyd, Walsh, & Shehni Yailagh, 2005), even when gender performance levels are the same. Thus, gender differences in math ability appear to be more a matter of self-perception than empirical confirmation.

Another factor that may temper the external validity of the current research was the fact that it was done with college students who should have had the mathematical skills to solve all the multiplication problems in all three assignments. In contrast to this assignment arrangement, homework assignments for school-age children are often given shortly after acquisition of a skill in order to produce mastery and fluency. Thus, children may react differently to assignments that target a newly developed rather than a well established skill, as was the case in the current study. Whether an assignment begins with longer or shorter problems may make a far greater difference if one has tenuous mastery of some tasks included in the assignment.

One over-riding premise behind the current study was that task order within assignments might affect students' assignment preference and performance, which in turn could affect whether a newly acquired skill becomes a mastered and fluent skill. One prospect is certain: a student must practice a newly developed skill in order to master it and become fluent in its use. If teachers can increase student practice of a target skill via in-class and homework assignments, they augment the likelihood that the child will master the skill, become fluent in the use of the skill, and generalize the skill to different situations. Thus, more research must be done to determine methods for increasing the likelihood of students' completing assignments, possibly by changing their perceptions of assignments, without reducing the quality or amount of practice embedded within assignments.

With the methodological changes previously suggested, task order may still prove valuable in altering student perceptions of the acceptability of assignments and choice of future assignments. Other research (e.g., Kern et al., 2001; Moes, 1998; Seybert et al., 1996; Watanabe & Sturmey, 2003) with individuals having psychological and behavioral problems shows that providing choices with regard to scheduling of assignments and the order of tasks within assignments can have a positive effect on an array of responses, including task engagement, completion rate, performance quality, and affective responses to assignments. Thus, continued research with general-education students, college level and below, may reveal some of the same benefits from allowing students to choose among task sequences that have been reported for assignment choices in research with students having special needs.

References

Balch, W. R. (1989). Item order affects performance on multiple-choice exams. *Teaching of Psychology, 16,* 75-77.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267-272.

Junge, M. E. & Dretzke, B. J. (1995). Mathematical self-efficacy gender differences in gifted/talented adolescents. *Gifted Child Quarterly, 39,* 22-28.

Kern, L., Mantegna, M. E., Vorndran, C. M., Bailin, D., & Hilt, A. (2001). Choice of task sequence to reduce problem behaviors. *Journal of Positive Behavioral Interventions, 3*(1), 3-10.

Laffitte, R. G. (1984). Effects of item order on achievement test scores and students' perception of test difficulty. *Teaching of Psychology, 11,* 212-213.

Lee, D. L. (2005). Increasing compliance: A quantitative synthesis of applied research on high-probability request sequences. *Exceptionality, 13,* 141-154.

Lee, D. L., Belfiore, P. J., Ferko, D., Hua, Y., Carranza, M., & Hildebrand, K., (2006). Using pre and post low-*p* latency to assess behavioral mometnum: A preliminary report. *Journal of Behavioral Education, 15,* 203-214.

Lloyd, J. E., Walsh, J., & Shehni Yailagh, M. (2005). Sex Differences in performance attributions, self-efficacy, and achievement in mathematics: If I'm so smart, why don't I know it? *Canadian Journal of Education, 28,* 384-408.

Moes, D. R. (1998). Integrating choice-making opportunities within teacher-assigned academic tasks to facilitate the performance of children with autism. *Journal of the Association of Persons with Severe Handicaps, 23,* 319-328.

Myers, D. L., & Myers, L. E. (1977). Undermatching: a reappraisal of performance on concurrent variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *25*, 203-214.

Neely, D. L., Springston, F. J., & McCann, S. J. H. (1994). Does item order affect performance on multiple-choice exams? *Teaching of Psychology, 21,* 44-45.

Nevin, J. A. (1992). An integrative model from the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior, 57,* 301-316.

Nevin, J. A. (1996). The momentum of compliance. *Journal of Applied Behavior Analysis, 29,* 535-547.

Nevin, J. A., Mandell, C., & Atak, J. R. (1983). The analysis of behavioral momentum. *Journal of the Experimental Analysis of Behavior, 39,* 49-59.

Olson, H. D., & Barickowski, R. S (1976). Adapting to multiple choice test of differing item arrangement. *Child Study Journal, 6,* 91-98.

Perlini, A. H., Lind, D. L., & Zumbo, B. D. (1998). Context effects on examinations: The effects of time, item order and item difficulty. *Canadian Psychology, 39,* 299-307.

Plaud, J. J., & & Gaither, G. A. (1996). Behavioral momentum: Implications and development from reinforcement theories. *Behavior Modification, 20,* 183-201.

Premack, D. (1959). Toward empirical behavior laws. *Psychological Review, 66,* 219-233.

Seybert, S., Dunlap, G., & Ferro, J. (1996). The effects of choice-making on the problem behaviors of high school students with intellectual disabilities. *Journal of Behavioral Education, 6*, 49-55.

Skinner, C. H. (2002). An empirical analysis of the interspersal research: Evidence, implications, and applications of the discrete task completion hypothesis. *Journal of School Psychology*, *40,* 347–368.

Watanabe, M., & Sturmey, P. (2003). The effect of choice-making opportunities during activity schedules on task engagement of adults with autism. *Journal of Autism and Developmental Disorders, 33,* 535-538.

Author Contact Information

Kathryn E. Jaspers
University of Tennessee
A525 Claxton Addition
1122 Volunteer Boulevard
Knoxville, TN 37996
kjaspers@utk.edu
865-974-4169

Christopher H. SKinner
University of Tennessee
A525 Claxton Addition
1122 Volunteer Boulevard
Knoxville, TN 37996
cskinne1@utk.edu
865-974-8403

Robert L. Williams
University of Tennessee
A525 Claxton Addition
1122 Volunteer Boulevard
Knoxville, TN 37996
bobwilliams@utk.edu
865-974-6625

Lee B. Saecker
University of Tennessee
A525 Claxton Addition
1122 Volunteer Boulevard
Knoxville, TN 37996
lsaecker@utk.edu
865-776-1822