Differentiating the Curriculum: What Difference Does it Make?

Roxane L. Lehmann
University or Minnesota
Paper prepared for AERA Annual Meeting 2002

It has been a persistent puzzle that students at both ends of the achievement spectrum often lack the motivational tendencies to allocate time and mental effort to learning tasks. That academically advanced students, identified for their potential to excel at school, often do not work diligently or persist when they encounter difficult learning tasks has been particularly puzzling (Ames, 1992; Bempechat, London, & Dweck, 1991; Dweck & Leggett, 1988; Schunk, 1994). Increasingly, theorists have acknowledged that the structures of the classroom environment affect students' motivational tendencies and learning outcomes (Ames, 1992; Urdan, 1997). Differentiating the curriculum by modifying the content, the process, and/or the product of the basic curriculum is one such classroom structure that has been widely recommended to respond to the diversity of academic needs typically found in mixed-ability middle school classrooms (e.g., Maker, 1989; Plucker & McIntire, 1996; Tomlinson, 1999). Developing differentiated curricula that provide multi-tiered learning activities to accommodate a variety of readiness levels has been suggested as an instructional framework to deal with academic diversity, particularly in meeting the needs of academically advanced students (Morelock & Morrison, 1997; Tomlinson, Moon, & Callahan, 1998). This type of instructional framework would allow students to make individually appropriate choices among tasks of graduated difficulty so that students are appropriately challenged and given more control over their learning.

Recent research on motivation within achievement goal theory, a theory concerned with students' purpose for achieving academically, supports the call of educational researchers for the use of differentiated curricula in mixed-ability classrooms. This research indicates that there is a connection between classroom structures, such as the design of the learning task and the task delivery process, and whether students' purpose for achieving is to demonstrate superior ability relative to others or to develop competence and mastery (Ames, 1992; Urdan, 1997). Further, students' purposes for achieving have been found to influence a range of motivational variables. Compared to students focused on demonstrating their ability, students focused on developing competence and mastery spend more time on learning tasks (Butler, 1987), are more persistent when confronted with a difficult task (Elliott & Dweck, 1988), and perform significantly better on the task (Mueller & Dweck, 1998).

Central to a focus on achieving to develop competence and mastery is a belief that effort and achievement outcome positively co-vary (Ames, 1992; Weiner, 1979). Differentiated curricula that embeds personal challenge in the design of learning tasks and opportunities for choice in the task delivery process can have positive effects on students' beliefs about the relation between effort and achievement outcome and their purpose for achieving. When undifferentiated curricula are used, tasks assigned by teachers are often poorly matched to students' skill level and abilities (Blumenfeld, 1992), and school performance and grades are generally determined far more by ability than effort (Schuman, et al., 1985). When differentiated curricula are used, tasks are better matched to students' skill levels and abilities. Further, the practice of differentiating the learning task should not be viewed as an independent contributor to students' beliefs about the relation between effort and achievement outcome. If the task structure focuses students' attention on effort in their explanations for their achievement outcome but the process of assigning tasks emphasizes social comparison information about ability, the positive consequences of task differentiation may be undermined. Providing multi-tiered learning activities to accommodate a variety of readiness levels and allowing for choice should make individual differences less conspicuous. Differentiated curricula coupled with task choice, thus, can help to provide students experiential evidence that fosters the belief that effort and achievement outcome positively co-vary. This belief can foster a focus on achieving to develop competence and mastery and result in beneficial effects on students' motivation (Ames; Malone & Lepper, 1987).

Although researchers in education and motivation have called for teachers to provide differentiated instruction in mixed-ability classrooms, most teachers practicing in the field do not agree with the researchers about the need for differentiation. The results of two national studies at the elementary and middle school level (Archambault, et al., 1993; Moon, Tomlinson, & Callahan, 1997) indicate most teachers believe there is no need to differentiate their curriculum. Findings from the current study conducted at a middle school provide evidence to help resolve this difference of opinion between researcher and practitioner.

Despite the substantial number of articles calling for differentiation and describing the structures of a differentiated classroom, little research exists on the effects of using a differentiated curriculum on the achievement behavior of academically advanced students or on students of other skill levels. This study, conducted using a randomized experimental design within a classroom context, was undertaken to provide empirical evidence that confirms or disconfirms the benefits of providing one type of differentiated instruction, multi-tiered learning tasks, particularly in responding to the needs of academically advanced students. Because providing opportunities for student choice in the structure of learning tasks is also considered an important factor in influencing students' motivation and quality of learning, this study took an integrative approach by considering the effects of the tasks' appropriateness as well as the task delivery process on achievement behavior.

The study examined the differential effects of the "one-size-fits-all" approach to instruction and the instructional practice of differentiating the structure of the task, so that the "task fits individual" in a teacher-assigned and a student-chosen condition on students of various skill levels. Specifically, the study measured the effects of the three conditions on (1) the co-variance of self-perceptions of challenge and time working on task, an index of effort, (2) the co-variance of time on task and score on task, and (3) a variety of achievement and reasoning behaviors for students of various skill levels. The achievement and reasoning behaviors examined included working minutes on task, self-perceptions of the task's challenge level, score on the task, and self-judgements about the quality of performance. It was predicted that when the task "fits the individual" in either the teacher-assigned task or student-chose task condition, students self-perceptions of the tasks' challenge would be more likely to co-vary with the time spent working on the task, and students would be more likely to earn scores that positively co-varied with the time spent on the task than in the control condition in which all students work on the same "one-size-fits-all" task of moderate difficulty. In addition, it was anticipated that students of high and low skill, the skill groups for whom the one-size-task-fits-all task is least appropriate, would be affected most by the differentiated task structure. It was predicted that when the assigned task was one that fit the individual rather than a "one-size-fits-all" task, (1) students of high skill would perceive the task as more challenging and work longer but their scores would be somewhat lower, and (2) students of low skill would perceive the task as less challenging, work longer, and earn higher scores. Finally, it was predicted that students at all skill levels would work longer, earn higher scores, and rate the quality of their performance as higher when a task that "fit the individual" was student-chosen rather than teacher-assigned.

Method

Participants. Sixth grade students (with an approximately equal number of girls and boys) were drawn from one public middle school in an upper-middle class, Midwestern suburb of the United States. All sixth graders within the school were invited to participate in the study.

Instruments. Two instruments were used prior to the experiment to obtain information about the students' scholastic profile. Students' self-perception of general academic competence was measured using The Scholastic Competence subscale of the Self-Perception Profile for Children (see Harter, 1985). In addition, students' achievement goal orientation was measured using an instrument that contains a choice of four task that embody different goals, i.e., performance or learning goals (see Dweck & Leggett, 1988; Mueller & Dweck, 1998).

Selected problems from Raven's Progressive Matrices (Raven, Court, & Raven, 1977, 1983) were used both to assess students' skills on the experimental task, so that blocks of students with similar ability could be established, and for the five forms of the experimental achievement tasks. Problems were selected based on item difficulty information derived from a pilot study done with approximately 100 fifth graders enrolled in an elementary school that feeds into the middle school where the research was conducted. The five forms of the experimental task varied in difficulty and length with more items included on the forms consisting of easier problems. The 20 items on the blocking task were representative of the range of difficulty of the items included on the five experimental task forms.

Two additional instruments were used as post-experimental measures. A single-item instrument developed for this study was used to assess the students' perceptions of the difficulty of the experimental task immediately after completing the task. The scale used to assess challenge judgements ranges from 10 to 100 in ten-unit intervals from low challenge (10) to high challenge (100). A second single-item instrument developed for this study was used to assess students' judgements about the quality of their performance, i.e., whether students view their performance on this task as a success or a failure. The item was scored from 1 to 4 with a score of 1 indicating very low perceptions of success and a score of 4 reflecting very high perceptions of success. The measure was administered immediately after students were shown their scores on the experimental task.

Procedures. This study included three sessions. The sessions were completed within a period of two weeks during the last term of the school year. All assessments were conducted by the researcher and four research assistants following standardized procedures.

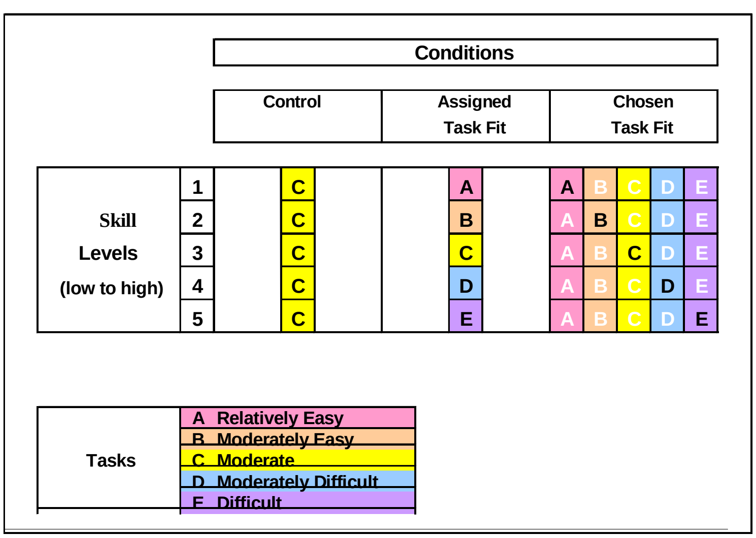
The blocking task used to assess each participant's skill level on the criterial task and the self-competence and achievement goal measures were administered to the students in one forty-five minute session in classroom groups. The skill assessment measure was administered first following a brief tutorial on solving Raven's puzzles. Students were encouraged to take as much time as they needed to complete the 20 puzzles.

Students were then classified according to skill in solving Raven's puzzles, the criterial task, and assigned to one of five skill blocks: low (8 or fewer correct), low-moderate (9 to 10 correct), moderate (11 to 13 correct), high-moderate (14 to 15 correct), and high (16 or more correct). One of three conditions was then randomly assigned to participants within each block. The conditions included a control in which students were assigned an undifferentiated, "one-size-fits-all" task of moderate difficulty (the type typically assigned to all students in a classroom) and two experimental conditions. In the first experimental condition, teacher-assigned "task fits individual," students were assigned a task differentiated to match their skill level using a multi-tiered task framework. In the second experimental condition, student-chosen "task fits individual," students were allowed to work at any one of the tiered tasks with analysis restricted to instances in which the chosen task matched the student's skill (i.e., the same task assigned to students within the corresponding block in teacher assigned "task fits individual" condition). (See Figure 1 below for a visual of the experimental design.)

To increase the potential for including sufficient numbers for analysis within the task choice condition, without drastically increasing the overall sample size, conditions were randomly assigned to participants in each block on a proportional basis. The "one-size-fits-all" control, the teacher-assigned "task-fits-individual," and the student-chosen "task-fits-individual" conditions were assigned to 25%, 25%, and 50% of the participants in each block, respectively.

The three conditions of the experiment were completed in a second thirty-minute group session within two days of the first session. In all conditions, participants were given their scores from the skill assessment measure. This feedback was provided in writing with the score reported as the number of problems solved correctly out of the total number (20).

Figure 1. Experimental Design



In the control condition, all participants were assigned the same task of "moderate" difficulty regardless of the participants' skill level on the task. Students were given feedback about the skill assessment completed in the first session and were told they would be given another puzzle task similar to the type worked in the first session. They were told that they would earn bonus points for doing this activity that could be exchanged for extra credit points on an upcoming assignment or test of their choice and that each problem answered correctly would be worth points. A strategy for solving the puzzles was briefly reviewed. Students then worked on the puzzle task. Immediately after finishing the task, students were asked to record the time at which they ended the task in a space provided on the answer sheet and to complete the single item designed to measure their perceptions of the difficulty of the puzzle task.

The procedures for the two experimental conditions were identical to the control procedures except for the manipulation of the independent "task" variable and the task description provided the participants. In the teacher assigned "task fits individual" condition the difficulty of the task varied by the participants' block assignment and was controlled by the researcher with the difficulty of the task assigned differentiated to match the participant's presumed skill based on the skill assessment measure. Table 1 below shows the scheme used to match the blocking task score with a task form.

Table 1. Assignments by Skill Block for "Assigned Task Matched Skill" Condition

Number Correct on Blocking Task	Skill Block	Description of Assigned Form
8 or fewer	Low	35 relatively easy puzzles
9 to 10	Low-Moderate	30 moderately easy puzzles
11 to 13	Moderate	25 moderate puzzles
14 to 15	High-Moderate	20 moderately difficult puzzles
16 to 20	High	15 relatively difficult puzzles

In this condition, after being told that the puzzles they would be given to work were similar to the type worked in the last session, students were told that not all of them would be working on the same puzzles and that not all of them would have the same number of puzzles. Instead, students were told that each of them would be working on a set of puzzles chosen specifically for them so that the task would be somewhat challenging but not too difficult for them to get many of them right. In addition, students were assured that all would have an equal chance to earn the same number of total bonus points. They were told that some tasks had more difficult puzzles than others, that more difficult puzzles were worth more points, and that tasks with more difficult puzzles had fewer puzzles than the easier tasks. The remainder of the session followed the procedures described above for the control condition.

In the student-chosen "task fits individual" condition, the difficulty of the task varied by participant and was controlled by the participant. In this experimental condition, after the students were told that the puzzles they would be given to work were similar to the type they worked in the last session, they were told that this time there were five different tasks made up of different puzzles and that they could choose which task to work. The same explanation given in the teacher-assigned "task fits individual" condition was used to assure students that regardless of the task they chose to work, they would all have an equal chance to earn the same number of bonus points. An overhead transparency was used to illustrate the differences among the five task forms described as: “35 relatively easy puzzles,” “30 moderately easy puzzles,” “25 moderate puzzles,” “20 moderately difficult puzzles,” and “15 relatively difficult puzzles.” In addition, students were provided a copy of the five tasks and given time to peruse these individually at their desks.

Students were encouraged to consider their scores earned on the skill assessment activity completed in session one when making their task selection. To help students judge the difficulty of the various tasks with their own skills, they were shown an overhead transparency depicting the strategy utilized in assigning experimental tasks to blocks of participants in the teacher-assigned "task fit individual” condition. The researcher said, “Try to choose a task that will be right for you. Think about choosing a task that will have puzzles that are somewhat challenging for you but not too difficult for you to get many of them right. In making your selection you may want to consider the score you earned on the set of puzzles you have done. This [overhead] shows one strategy you might use to choose which set of puzzles to work. This is one way that you might match your skill at solving these puzzles with the task that is right for you. This is a strategy someone might use if the only information they had about you was the score on the puzzles you did earlier. This score, however, is just one of many pieces of information you have about yourself as a learner.” Then, following the same procedure used in the first session to ensure that participants felt no demand from the researcher to select a particular task, participants were told, “Some kids like to choose different tasks. You should pick the task that you think is right for you. I am interested in what task you will choose.” Students were given several minutes to select a task. Once task selections had been made, the remainder of the session followed the procedures described above for the control condition.

In a subsequent ten-minute session, within two days of the second session, students were shown their scores on the experimental task and questioned regarding their post-performance judgements about the quality of their performance. For this session, students were seen individually by one of four female experimenters, blind to both the experimental hypotheses and the condition of each participant.

Results

Description of the sample. Of the 204 students recruited for this study, a total of 186 (99 girls and 87 boys) completed all three of the study's sessions. Of these 186 students, 91 were assigned to the condition in which they were allowed to work at the task of their choice. Only the 38 of these students who chose the task that aligned with their identified skill level, 42% of the students assigned to this condition, were included in the study's analyses as participants in the student-chosen "task fits individual" condition. The remaining 53 students assigned to this condition were excluded from the analyses. In addition, two participants (in the teacher-assigned "task fits individual" condition) were excluded from analyses because their scores on the criterial task were extremely low (less than 40% correct). The study's analyses, thus, included the 47 participants assigned to the control condition, the 46 participants in the teacher-assigned "task fits individual" condition, and the 38 participants in the student-chosen "task fits individual" condition.

Preliminary analysis. A series of one-way analyses of variance (ANOVAs) was conducted for each skill block to verify that the 38 students in the student-chosen "task fits individual" condition who selected the task matched to their identified skill level were similar to students assigned to the other two conditions of the experiment. This analysis examined the demographic characteristics, and scholastic profile variables (i.e., self-perceptions of scholastic competence, achievement goal orientation, score earned on the blocking task, and time spent on the blocking task) of students across the three conditions. Only two effects were found and neither affected the interpretation of the study's findings. An additional series of one-way ANOVAs was conducted to examine the effects of different experimenters, teaching teams, and student gender on responses to the dependent measures. No effects were found.

Further analyses were conducted to insure that that the blocking task of 20 problems was a valid measure of the participants' skill at solving the puzzles used to form the experimental tasks. The correlation between the blocking task scores and experimental scores for the control group, in which all students worked on the same task of “moderate” difficulty, was found to be positive and significant, r = .562, p Ω .001. In addition, a one-way ANOVA revealed that scores on the experimental task in the control condition differed according to the skill block assignment, F (2, 38) = 12.96, p Ω .001.

Finally, in analyses of the dependent measures conducted to determine whether assumptions for analysis of variance were satisfied, two violations to the homogeneity of variance assumption were found. This heterogeneity in variance, found in analyses of the data for scores on the experimental task and working minutes on the experimental task, was anticipated due to the difference in task difficulty between the control condition and the two matched conditions. Because the largest variances for the task score data occurred in the groups with the largest number of participants, the test statistic will be conservative (Stevens, 1986), and no adjustments or modifications to the data were made. The largest variances for the working minutes on task data, however, did not occur in the groups with the largest number of participants. Therefore, the minutes on task data were transformed using log base 10. After modification, the data fit more closely the underlying assumptions for ANOVA with the largest variance less than four times the smallest, and the transformed data was used for analyses.

Effects of the experimental conditions on the relations between working minutes, challenge perceptions, and score. Means and standard deviations for working minutes on task, students' self-perceptions of challenge, and score on the task for the three experimental conditions are shown in Table 2. The minutes spent working on the tasks across conditions are in line with expectations with each form of the task designed to take 10 to 15 minutes to complete. In addition, although there was significant variation in students' perceptions of the difficulty of the tasks, on average students across the conditions perceived the task to be within the "somewhat challenging" range. Finally, the average scores across the three conditions indicate that overall students did well on the puzzle tasks.

A partial correlation controlling for student's perception of challenge, a moderator variable, revealed students earned scores that positively co-varied with time spent working on the task only in the student-chosen "task fits individual" condition in which students chose the task matched to their skill level ($r = .35$, $p \leq .05$). No significant relation between working minutes was found for either the "one-size-fits-all" control group in which an undifferentiated task was assigned (r = 14, ns) or the teacher-assigned "task fits individual" experimental group in which a task matched to the student's skill level was assigned (r = -.09, ns).

Table 2. Means and Standard Deviations for Working Minutes on Task, Self-Perceptions of Challenge, and Score (% correct)

Dependent Variable		Experimental Condition		
		Control	Assigned Fit	Choice Fit
Working Minutes				
	Mean	8.40	9.48	9.66
	SD	2.13	3.16	3.34
	N	43	46	35
Self-Perceptions of Challenge				
	Mean	42.87	45.32	39.61
	SD	26.58	23.58	25.79
	N	47	47	38
Score				
	Mean	84.94	79.11	82.11
	SD	13.22	13.98	13.01
	N	47	47	38

In addition, bivariate correlations revealed a positive association between students' self perceptions of the challenge level of the task and the time they spent working on the task only for participants in the student-chosen "task fits individual" condition (r = .40, p ≤ .01). Interestingly, this relation also was found to be significant when the data for all 91 students assigned to this condition, regardless of the appropriateness of the task chosen, were analyzed (r = .35, p ≤ .001). No significant relation between perceptions of challenge and time spent on task was found for either the control group (r = .14, ns) or the teacher-assigned "task fits individual" group (r = .19, ns).

Effects of the experimental conditions on students of different skill levels. As described previously under methods, participants were divided into five groups based on their scores on the first set of problems, an index of their ability on the task. Of the initial 131 students in the combined "one-size-fits-all" control, the teacher-assigned "task fits individual," and the student-chosen "task fits individual" conditions, only four were in the low skill block and eight were in the low-moderate skill block. Students in these skill blocks were excluded from the two-way analyses of variance (ANOVAs) because these blocks contained too few participants to include in the analyses conducted to examine the effect of the experimental conditions across the skill blocks. In summary, the ANOVA analyses included three skill blocks (moderate, high-moderate, and high skill) with 41 participants assigned to the control condition, 43 participants in the teacher-assigned "task fits individual" condition, and 35 participants in the student-chosen "task fits individual" condition.

After running a multivariate analysis to control for study-wise alpha (Pillai's Trace = .668 p ≤ .0005), two-way ANOVAs (criterial task condition by students'skill block) were conducted on the dependent measures of working minutes on task, self-perceptions of the task's challenge level, score on task, and self-judgements about the quality of performance. The results of these ANOVAs are summarized in Table 3. Simple effects were analyzed by skill level when significant interaction effects were found. When significant simple effects were found, tests of multiple comparisons were run to further tease apart differences. Means and standard deviations as well as the interaction effects from the overall analyses and the follow-up simple effects of the dependent variables for the three sub-sample skill levels across the three research conditions are shown in Tables 4 and 5.

As depicted in Figure 2a**,** the effects of the condition on working minutes depended on skill level. Specifically, the analysis revealed an interaction of condition by skill, F (4, 102) = 2.76, p Ω .05. The within-skill-levels analyses of simple effects indicated that whereas students of moderate and high-moderate skill did not vary in working minutes on task across conditions, the time spent on task did vary across conditions for the highly skilled students, F (2, 102) = 3.69, p Ω .05. The students with high skill at the experimental task worked longer in the two conditions in which the task was aligned with their skill level than in the control condition. Tests of multiple comparisons run for the high skilled group indicated the difference in working minutes between the control (M = 8.47, SD = 1.74) and the student-chosen "task fit individual condition (M = 11.69, SD = 3.99) was significant,

Table 3. ANOVA Source Tables for Experimental Task Working Minutes, Self-Perceptions of Challenge, Score, and Judgements about the Quality of Performance (Skill Block by Experimental Condition Models)

Variable	Source	*df*	SS	MS	F
Working Minutes $_a$					
	Block	2	.05	.02	1.69
	Condition	2	.05	.03	1.72
	Block X Condition	4	.16	.04	2.76*
	Error	102	1.50	.01	
Self-Perceptions of Challenge					
	Block	2	1491.87	745.94	1.45
	Condition	2	820.48	410.24	.80
	Block X Condition	4	14248.96	3562.24	6.94***
	Error	110	56444.83	513.14	
Score					
	Block	2	1414.28	707.14	6.14**
	Condition	2	614.05	307.02	2.67
	Block X Condition	4	6376.09	1594.02	13.85***
	Error	110	12661.15	115.10	
Performance Judgements					
	Block	2	.88	.44	.82
	Condition	2	.74	.37	.69
	Block X Condition	4	13.23	3.31	6.13***
	Error	110	59.33	.54	

a. Data transformed using Log Base 10
*p. ≤ .05 **p. ≤ .01 ***p. ≤ .001

p Ω .05, with an effect size of .88, whereas the difference in working minutes between the control and teacher-assigned "task fit individual" condition (M = 9.33, SD = 2.06) was not significant.

The analysis of perceptions of challenge revealed an interaction of condition by skill level, F (4, 110) = 6.94, p Ω .001. The analysis of simple effects indicated perceptions of challenge varied across conditions for students at all three skill levels: moderate, F (2,110) = 3.34, p Ω .05; high-moderate, F (2,110) = 3.36, p Ω .05; and high, F (2,110) = 7.57, p Ω .001. As figure 2b shows, in the control and teacher-assigned "task fits individual" conditions, students of moderate (M = 56.00, SD = 17.76 and M = 52.31, SD = 25.22, respectively) as well as those of high-moderate skill (M = 44.29, SD = 27.86,

Table 4. Simple Effects for Experimental Task Working Minutes and Self-Perceptions of Challenge

Dependent Variable	Skill Block	Experimental Condition			Interaction Coefficient F	Simple Effect F
		Control	Assigned Task Fit	Chosen Task Fit		
Working Minutes					2.76*$_a$	
	Moderate					1.36$_a$
	Mean	9.00	8.23	9.89		
	SD	2.62	2.24	2.09		
	N	8	13	9		
	High-Moderate					2.99$_a$
	Mean	8.25	10.07	7.90		
	SD	2.45	2.67	1.79		
	N	12	14	10		
	High					3.69$_a$*
	Mean	8.47	9.33	11.69		
	SD	1.74	2.06	3.99		
	N	17	15	13		
Self-Perceptions of Challenge					6.94***	
	Moderate					3.34*
	Mean	56.00	52.31	32.00		
	SD	17.76	25.22	19.32		
	N	10	13	10		
	High-Moderate					3.36*
	Mean	44.29	46.43	24.55		
	SD	27.86	24.68	15.08		
	N	14	14	11		
	High					7.57***
	Mean	32.06	40.63	63.21		
	SD	25.19	21.12	20.53		
	N	17	16	14		

a. Data transformed using log Base 10
*p. ≤ .05 **p. ≤ .01 ***p. ≤ .001

and M = 46.43, SD = 24.68, respectively) perceived the task as equally challenging. Both moderate (M = 32.00, SD = 19.32) and high-moderate (M = 24.55, SD = 15.08) groups, however, perceived the task as less challenging in the student-chosen task fits individual condition. Tests of multiple comparisons on these simple effects provided no additional information about differences in perceptions of challenge among the conditions for either of these skill blocks. In contrast, students of high skill perceived their task as more challenging in the teacher-assigned "task fit individual" condition (M = 40.63, SD = 21.12) than in the control condition (M = 32.06, SD = 25.19) and most challenging in the student-chosen "task fits individual" condition (M = 63.21, SD = 20.53). Tests of multiple

Table 5. Simple Effects for Experimental Task Scores and Judgements about the Quality of Performance

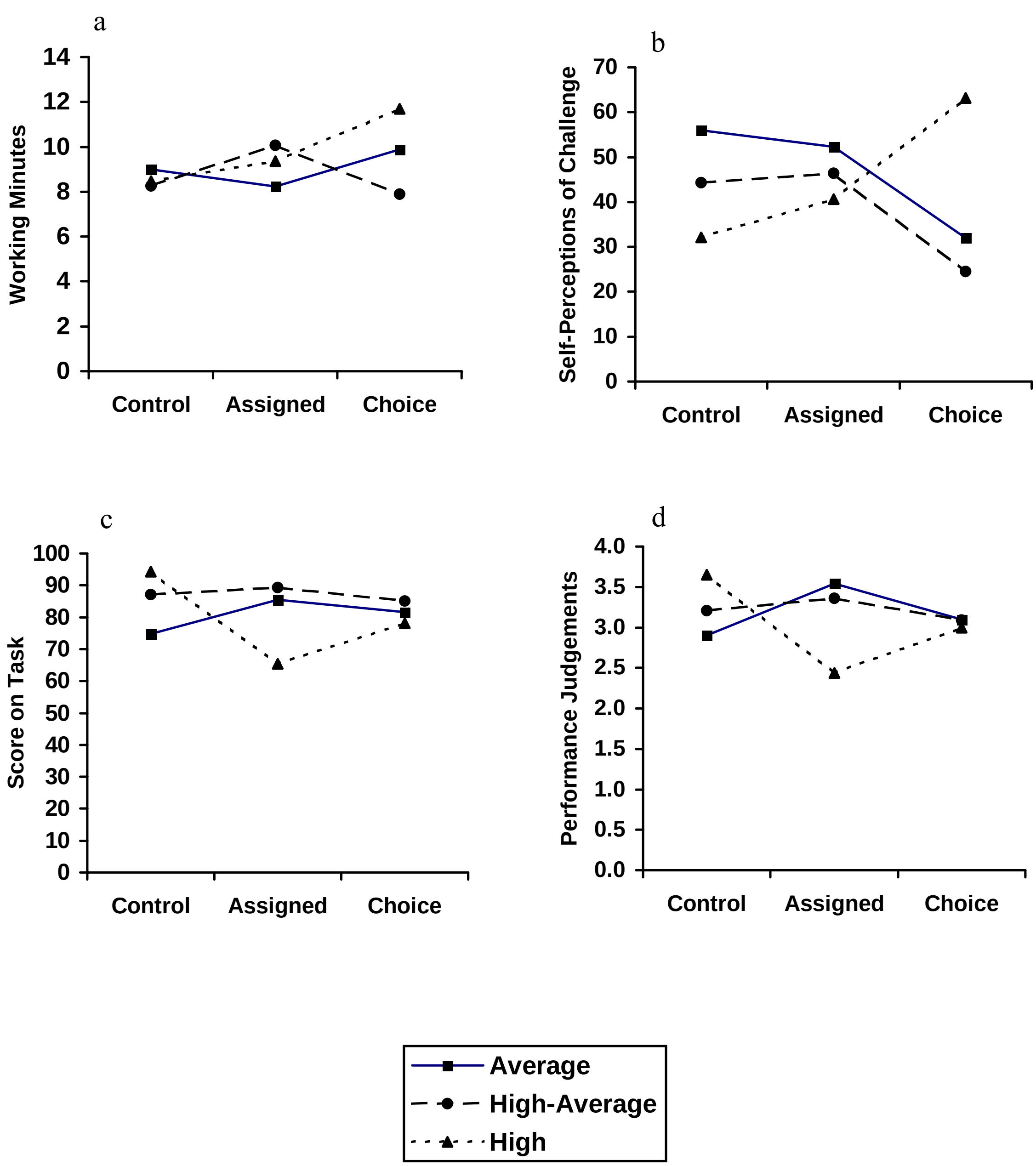
Dependent Variable	Skill Block	Experimental Condition			Interaction Coefficient F	Simple Effect F
		Control	Assigned Task Fit	Chosen Task Fit		
Score					13.85***	
	Moderate					2.85
	Mean	74.80	85.54	81.60		
	SD	13.99	8.57	13.49		
	N	10	13	10		
	High-Moderate					.49
	Mean	87.14	89.29	85.00		
	SD	9.94	9.17	10.25		
	N	14	14	11		
	High					30.27***
	Mean	94.35	65.38	78.14		
	SD	5.49	10.09	15.05		
	N	17	16	14		
Performance Judgements						
	Moderate					2.30
	Mean	2.90	3.54	3.10		
	SD	.57	.66	.74		
	N	10	13	10		
	High-Moderate					.41
	Mean	3.21	3.36	3.09		
	SD	.80	.63	.94		
	N	14	14	11		
	High					11.22 ***
	Mean	3.65	2.44	3.00		
	SD	.49	.81	.88		
	N	17	16	14		

***p. ≤ .001

comparisons indicated that perceptions of challenge for the high skill students in the student-chosen "task fits individual" condition differed significantly from the perceptions of students in the control (p Ω .001) and in the teacher-assigned "task fits individual"

For the analysis of scores on the experimental task, there was an interactions of skill level by experimental condition, F (4, 110) = 13.85, p Ω .001. The analysis of simple effects conducted to help interpret the interaction effect, showed that whereas scores for the moderate and high-moderate students were not significantly different across conditions,

Figure 2. Effects of Experimental Conditions on Students of Different Skill Levels



Control Assigned Choice

scores for the students of high skill did vary significantly, F (2, 110) = 30.27, p Ω .001. Tests of multiple comparisons indicated the differences in scores for these students were significant among all three conditions. Again, because the task for the high-skilled students was significantly less difficult in the control condition, specific group differences between the control and the two experimental groups were expected. However, the significant difference in scores found for high skilled students between the two experimental groups, teacher-assigned "task fit individual" (M = 65.38, SD = 10.09) and student-chosen "task fit individual" (M = 78.14, SD = 15.05), p Ω .05, illustrated in Figure 2c, and the large effect size of 1.01 are noteworthy.

Finally, analysis of student judgements about the quality of their performance indicated there was an interaction effect, F (4,110) = 6.13, p Ω .001, depicted in Figure 2d. Simple effects analyses showed that whereas the performance judgements of students of moderate and high-moderate skill were not significantly different across conditions, the performance judgements of the highly skilled students did vary significantly, F (2, 110) = 11.22, p Ω .001. Although it was assumed that the significantly higher scores earned by students in the high skill block in the control condition would generate higher performance judgements than those of students with similar skill in either experimental condition, tests of multiple comparison did not confirm this assumption. Instead, results of tests of multiple comparisons indicated that whereas performance judgements for high skilled students were significantly lower in the teacher-assigned "task fit individual" condition (M = 2.44, SD = .81) than in the control condition (M = 3.65, SD = .49), p Ω .001, the performance judgements of these students in the control condition and the student-chosen "task fits individual" condition (M = 3.00, SD = .88) were not significantly different. The computed effect size indicated that the mean of the students' performance judgements in the control condition was 1.83 standard deviations above that of the teacher-assigned "task fit individual" condition.

Discussion

This study examined the potential benefits of providing one type of differentiated instruction, multi-tiered learning tasks, particularly in responding to the needs of academically advanced learners. The study took an integrative approach, examining both the design of the task and the task delivery process on students' reasoning and achievement behavior. The study examined the differential effects of a "one-size-fits-all" approach to instruction and instruction that differentiates the curriculum to fit the individual student using multi-tiered learning tasks that were either assigned to the student or chosen by the student. The primary goals of the study were to find out (1) whether the co-variance of time spent and score on the task would differ across the conditions and (2) whether students' learning behavior and perceptions about the task and their performance would differ across the conditions. The results of this study provide clear evidence of the differential effects that "one-size-fits-all" tasks and student-chosen tasks that fit the individual have on students' reasoning and achievement behavior. Although the teacher-assigned tasks that fit the individual did not have the predicted beneficial effects on achievement behavior, there were clear benefits when students chose tasks that fit their individual skill levels.

Overall, this study's findings indicate that students benefited when personal challenge was built into the structure of the learning task and choice was built into the task delivery process. Across skill levels, time spent working on the task and the score earned on the task positively co-varied only for students who selected an appropriate task so that the task fit their individual skill level.

But why would time and score co-vary when students chose a task that fit their skill level and not when that same task was assigned? Before choosing which task to work, participants in the student-chosen "task fits individual" condition were given an opportunity to peruse the multi-tiered tasks and were provided information about the relative difficulty of each task and how to judge the difficulty of the various tasks with their own skill. Apart from the task delivery process, this opportunity to compare and contrast the tasks and access to a strategy for selecting a task that fit their skill were the only variations between the two experimental "task fit individual" conditions. Perhaps these provisions convinced the participants in the student-chosen "task fits individual" condition, in a way that the those in the teacher-assigned condition were not, that the task's challenge was appropriate and that additional time spent working would be worth the effort. Regardless of the underlying causes, when students chose a task that fit their skill level, their perceptions of the task's challenge level predicted the time they spent working on the task. The more difficult they perceived the task, the longer they worked, and those who worked longer on the task earned higher scores.

Further, for students of high skill who chose the task that fit their skill level, rather than having that same task assigned, there were additional benefits. Highly skilled students who were allowed to choose their task, perceived the task as significantly more challenging than students of the same skill level who were assigned an identical task. In addition, the highly skilled students who chose their task earned significantly higher scores (13% points higher) than similar students assigned the same task. Being allowed to select the task also positively affected the amount of time these students spent working on the task and their judgements about the quality of their performance.

Similar positive effects were not observed for those of high-moderate or moderate skill, although it was predicted that students at all skill levels would benefit from task choice. It is possible that the differential impact of task choice across skill levels may be related to the documented difference in students' perceptions of their tasks' challenge level within skill levels between the two experimental conditions. In contrast to the heightened perceptions of challenge observed in highly skilled students who chose their task, students of high-moderate and moderate skill who were allowed to choose their task perceived the task as significantly less challenging than students of the same skill level who were assigned an identical task. These differences may be related to the opportunity to peruse all the tasks and by information, such as the relative difficulty of each task, within the task selection strategy provided participants in the student-chosen "task fits individual" condition. Students of high skill in this condition who selected to work the most difficult task knew that their chosen task was designed to be the most challenging. Similarly, students of moderate and high-moderate skill who chose to work the task designed to fit their skill level knew that their chosen task was not the most challenging task available. In contrast, students in the teacher-assigned "task fits individual" condition were given no supplemental information about the relative difficulty of their task. Alternatively, it is possible that the high-skilled students simply benefited more from the sense of autonomy and control afforded participants in the student-chosen "task fits individual" condition than did students of lesser skill. High achievers tend to prefer a greater degree of control than average students in determining the activities they will pursue (Whitmore cited in Middleton, Littlefield &Lehrer, 1992) and may benefit more from having task options open to them (Marshall & Weinstein, 1984). Finally, it is important to note that although students of lesser skill apparently did not benefit from task choice as did highly skilled students, results suggest that they were not negatively affected either. Although students of moderate and high-moderate skill perceived the task as less challenging, differences in working time, scores earned, and self-judgements about the quality of performance for these students were not significant.

Conclusion

The results of this study underscore that classroom structures should not be viewed as independent contributors to students' achievement behavior. Providing teacher-assigned tasks that were differentiated to fit the individual did not enhance students' motivational tendencies to allocate time and mental effort to learning tasks. Instead beneficial changes in students' achievement behavior were noted only when providing a task that fit the individual was paired with task choice. The findings, thus, indicate that these instructional practices are interdependent and should be coordinated so that they work in concert.

The need to coordinate these instructional practices generates new questions that need to be addressed to support teachers in their efforts to optimally respond to the diverse needs within a mixed-ability classroom. The beneficial effects of allowing students to choose among multi-tiered task options paired with the need for these students to choose an appropriate task leads to questions about how teachers can provide students with both support and autonomy in choosing among task options. Linked with this question are the related questions about what choices students perceive as real and attractive options and how to ensure that students perceive the tasks as equal so that their choices are not determined by an interest in minimizing work.

Teachers who have differentiated their curriculum often cite a concern that students will simply choose the easiest task as a rationale for not allowing students to choose among differentiated task options. Yet in this study 42% of the students chose the task that was designed to fit their skill level and a full 75% of them chose a task that was within one level of the task that was considered most appropriate for their skill. This leads to another set of questions about what factors need to be embedded in the structure of the learning tasks and the task delivery process to encourage students to choose appropriately.

Finally, it is important to recognize that the findings may reflect features of the particular context in which the study was conducted. All of the participants were from a middle school with a higher than average ability student population. In addition, the tasks were novel rather than actual school tasks, and scores earned translated into bonus points rather than grades. Future research should address questions about the degree to which the experimental findings can be replicated in schools with a more diverse student population, including a larger number of students with less than average skills, using actual school tasks, and a more typical method of grading.

Although questions remain to be investigated, the results of this research indicate that the classroom structures of task differentiation and task choice that are theoretically related to student achievement behavior are related in practice as well. This study was conducted within a classroom context, and the structure of the learning tasks and the task delivery process used within this study could be applied in any classroom. The results indicate that embedding personal challenge in the design of the learning task and opportunities, as well as support, for choice in the structure of the task delivery process can have beneficial effects on student motivation and achievement outcome. Differentiating the curriculum by providing multi-tiered learning tasks so that students work on an appropriately challenging task can help to provide students with experiential evidence that effort positively co-varies with achievement outcome. Further, providing students with support and autonomy so that they make appropriate choices among the multi-tiered learning task options can benefit highly skilled students without negatively affecting students of lesser skill. Coordinating these structures so that they work in concert to meet the diverse needs within a mixed-ability classroom could make students' responses to classroom learning situations much less puzzling.

References

Ames, C. (1992). Classroom: Goals, structures, and student motivation. Journal of Educational Psychology, 84 (3), 261-271.

Archambault, F. X. Jr., Westberg, K. L., Brown, S. W., Hallmark, B. W., Emmons, C. L., & Zhang, W. (1993). Regular classroom practices with gifted students: Results of a national survey of classroom teachers. University of Conneticut, Storrs, CT: The National Research Center on the Gifted and Talented (NRC/GT).

Bempechat, J., London, P., & Dweck, C. S. (1991). Children's conceptions of ability in major domains: An interview and experimental study. Child Study Journal, 2 (1), 11-36.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. Journal of Educational Psychology, 79 (4), 474-482.

Dweck, C. S., & Leggette, E. L. (1988). A social-cognitive approach to motivation and personality. Psychological Review, 95 (2), 256-273.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. Journal of Personality and Social Psychology, 54 (1), 5-12.

Harter, S. (1985). Manual for the self-perception profile for children. University of Denver.

Maker, C.J. (1989). Curriculum content for gifted students: Principles and practices. In R. M. Milgram (Ed.), Teaching gifted and talented learners in regular classrooms (pp. 33-61). Springfield, IL: Charles C Thomas Publisher.

Malone, T. W., & Lepper, M.R. (1987). Making learning fun: A taxonomy of intrinsic motivation for learning. In R.E. Snow & M.J. Farr (Eds.), Aptitude, learning, and instruction: Vol. 3 (pp. 255-286). Hillsdale, NJ: Erlbaum.

Marshall, H. H., & Weinstein, R. S. (1984). Classroom factors affecting students' self-evaluations: An interactional model. Review of Educational Research, 54 (3), 301-325.

Middleton, J. A., Littlefield, J., & Lehrer, R. (1992). Gifted students' conceptions about academic fun: An examination of a critical construct for gifted education. Gifted Child Quarterly, 36 (1), 38-44.

Moon, T., Tomlinson, C. A., & Callahan C. M. (1997). Academic diversity in the middle school: Results of a national survey of middle school administrators and teachers. University of Connecticut, Storrs, CT: The National Research Center on the Gifted and Talented (NRC/GT).

Morelock, J., & Morrison, K. (1997). Differentiating 'developmentally appropriate': The multidimensional curriculum model for young gifted children. Roeper Review, 21 (3), 195-200.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. Journal of Psychology and Social Psychology, 75 (1), 33-52.

Plucker, J. A. & McIntire, J. (1996). Academic survivability in high-potential, middle school students. Gifted Child Quarterly, 40 (1), 7-14.

Raven, J. C., Court, J. H., & Raven, J. (1977). Advanced progressive matrices. London: H. K. Lewis & Co.

Raven, J. C., Court, J. H., & Raven, J. (1983). Standard progressive matrices. London: H. K. Lewis & Co.

Schuman, H., Walsh, E., Olson, C., & Etheridge, B. (1985). Effort and reward: The assumption that college grades are affected by quantity of study. Social Forces, 63 (4), 945-966.

Schunk, D. H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D. H. Schunk & B. J. Zimmerman (Eds.), Self-regulation of learning and performance (pp. 75-99). Hillsdale, NJ: Erlbaum.

Stevens, J. (1986). Applied multivariate statistics for the social sciences (pp. 216-226). Hillsdale, NJ: Erlbaum. Stevens, J. (1986). Applied multivariate statistics for the social sciences (pp. 216-226). Hillsdale, NJ: Erlbaum.
Tomlinson, C. A. (1999). The differentiated classroom: Responding to the needs of all learners. Alexandria, VA: ASCD.

Tomlinson, C. A., Moon, T. R., & Callahan, C. M. (1998). How well are we addressing academic diversity in the middle school? Middle School Journal, 29 (3), 3-11.

Urdan, T. C. (1997). Achievement goal theory: Past results, future directions. Advances in Motivation and Achievement, Vol. 10, (pp. 99-141). Greenwich, CT: JAI Press.

Weiner, B. (1979). A theory of motivation for some classroom experiences. Journal of Educational Psychology, 71, 3-25.