International Journal of Educational Research 50 (2011) 117–136

Contents lists available at ScienceDirect

International Journal of Educational Research

journal homepage: www.elsevier.com/locate/ijedures

Longitudinal investigation of the curricular effect: An analysis of student learning outcomes from the LieCal Project in the United States

Jinfa Cai [a,*], Ning Wang [b], John C. Moyer [c], Chuang Wang [d], Bikai Nie [a]

[a] University of Delaware, Newark, USA
[b] Widener University, Chester, USA
[c] Marquette University, Milwaukee, USA
[d] University of North Carolina – Charlotte, Charlotte, USA

ARTICLE INFO

Keywords:
Longitudinal study
Mathematics curriculum reform
Learning of algebra
Growth curve modeling

ABSTRACT

In this article, we present the results from a longitudinal examination of the impact of a *Standards*-based or reform mathematics curriculum (called CMP) and traditional mathematics curricula (called non-CMP) on students' learning of algebra using various outcome measures. Findings include the following: (1) students did not sacrifice basic mathematical skills if they are taught using a *Standards*-based or reform mathematics curriculum like CMP; (2) African American students experienced greater gain in symbol manipulation when they used a traditional curriculum; (3) the use of either the CMP or a non-CMP curriculum improved the mathematics achievement of all students, including students of color; (4) the use of CMP contributed to significantly higher problem-solving growth for all ethnic groups; and (5) a high level of conceptual emphasis in a classroom improved the students' ability to represent problem situations. (However, the level of conceptual emphasis bears no relation to students' problem solving or symbol manipulation skills.)

Published by Elsevier Ltd.

The ultimate goal of educational reform, curriculum innovation, and instructional intervention is to improve students' learning. By means of changes in curricula, advocates of mathematics education reform often attempt to change classroom practice, and hence, students' learning (Ball & Cohen, 1996; Clements & Sarama, 2008; Howson, Keitel, & Kilpatrick, 1981; NCTM, 1989; Senk & Thompson, 2003).

In the third article of this special issue (see Moyer et al.), we documented the impact of curriculum innovation on classroom instruction. In particular, using data from our project, Longitudinal Investigation of the Effect of Curriculum on Algebra Learning (LieCal), we examined the similarities and differences between a *Standards*-based curriculum, called the Connected Mathematics Program (CMP), and more traditional curricula, called non-CMP curricula. As we have pointed out, the LieCal Project was designed to compare longitudinally the effects of the CMP curriculum to the effects of more traditional middle school curricula on students' learning of algebra. In the LieCal Project, we investigated not only the ways and circumstances under which the CMP and non-CMP curricula promoted or hindered student achievement gains, but also the characteristics of the reform and traditional curricula that contribute to these gains. In this article, we provide evidence of the effect of these curricula on student learning.

What really works for improving students' learning then? This is the most frequently asked question in the current debate about the mathematics education reform movement in the United States, particularly with regard to the recent *Standards*-base curriculum innovations (e.g., Herman, Boruch, Powell, Fleischman, & Maynard, 2006; Schoenfeld, 2006). It has been

* Corresponding author at: 523 Ewing Hall, University of Delaware, Newark, DE 19716, USA.
E-mail address: jcai@udel.edu (J. Cai).

0883-0355/$ – see front matter. Published by Elsevier Ltd.
doi:10.1016/j.ijer.2011.06.006

claimed that the main emphasis of the *Standards*-based curricula is on conceptual understanding through problem solving, rather than on procedural knowledge. Supporters of reform maintain that students will learn procedural knowledge and master basic skills as they engage in explorations of worthwhile problems (NCTM, 2000). Nonetheless, many parents and teachers worry that if *Standards*-based curriculum innovations are implemented, any potential development of students' higher-order thinking skills will come at the expense of the development of basic mathematical skills (e.g., Cai, 2003; Wu, 1997). Therefore, an over-riding question about curriculum reform is *Do conceptual understanding and higher order thinking skills come at the expense of basic mathematics skills for students who are taught using a Standards-based mathematics curriculum?*

In the United States, classrooms are becoming more ethnically diverse. Since teaching and learning are cultural activities, students with different ethnic and cultural backgrounds may respond differently to the same curriculum. Therefore, a second over-riding question about curriculum reform is How does *the use of a Standards-based curriculum impact the learning of students of color as compared to Caucasian students*? In this article, we provide evidence of the effect of curriculum reform on student learning by addressing these two questions.

1. Theoretical considerations

1.1. Background

In the United States, the National Council of Teachers of Mathematics (NCTM) has provided recommendations for reforming and improving K-12 school mathematics through its *Standards* documents (1989, 2000). In these and related documents, the discussions of goals for mathematics education focus on the importance of thinking, understanding, reasoning, and problem solving, with an emphasis on connections, applications, and communication. This view stands in contrast to a more conventional view of mathematics education, which involves the memorization and recitation of de-contextualized facts, rules, and procedures, with an emphasis on the application of well-rehearsed procedures to solve routine problems.

To make curricula that align with the NCTM standards available to teachers, the U.S. National Science Foundation provided support to develop a number of so called *Standards*-based school mathematics curricula. (See Senk and Thompson (2003) for information regarding these NSF-funded curricula.) CMP is one of these *Standards*-based curricula. It is a complete middle-school mathematics curriculum that was identified as an exemplar by the U.S. Department of Education (U.S. Department of Education, 1999). The intent of CMP is to build students' understanding of major ideas in number, algebra, geometry, measurement, data analysis, and probability through explorations of real-world situations and problems (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2002a). NSF-funded curricula like CMP not only look very different from commercially developed (i.e., traditional) mathematics curricula, but they also have different kinds of learning goals.

Field tests of *Standards*-based middle school curricula have shown that on standardized tests measuring computational skills and procedural knowledge, students using *Standards*-based curricula performed at least as well as students using traditional curricula (Senk & Thompson, 2003). In addition, they have shown that students using *Standards*-based curricula performed better than students using traditional curricula on tests specifically designed to measure conceptual understanding and problem solving. By the latter part of the 1990s, school districts across the United States began to formally adopt the *Standards*-based mathematics curricula. Since then, a few studies have investigated their effect on the acquisition of basic skills and procedural knowledge (Harwell et al., 2007; Post et al., 2008; Reys, Reys, Lappan, Holliday, & Wasman, 2003; Riordan & Noyce, 2001). These studies generally have shown that neither students' facility with basic skills nor their procedural knowledge was harmed by using *Standards*-based curricula.

These findings are certainly significant. However, to date, in districts that have formally adopted *Standards*-based curricula, there have been no comprehensive longitudinal studies of the effect of *Standards*-based curricula on students' learning. In particular, except for our LieCal Project, there have been no long-term longitudinal studies of the effect of *Standards*-based curricula on students' learning of algebra.

1.2. LieCal project

The CMP curriculum was selected for investigation in the LieCal Project for several reasons, not the least of which is the fact that it has been more broadly implemented than any other *Standards*-based curriculum at the middle school level. In the 2002–2003 school year, CMP was used in nearly 2500 school districts in the United States. It has been used in all 50 states and some foreign countries (Rivette, Grant, Ludema, & Rickard, 2003; Show-Me Center, 2002).

By comparing, longitudinally, the effects of the CMP curriculum on students' learning of algebra to the effects of more traditional middle-school mathematics curricula (hereafter called non-CMP curricula), the LieCal Project is designed to provide: (a) a profile of the intended treatment of algebra in the CMP curriculum with a contrasting profile of the intended treatment of algebra in non-CMP curricula; (b) a profile of classroom experiences that CMP students and teachers have, with a contrasting profile of experiences in non-CMP classrooms; and (c) a profile of student performance resulting from the use of the CMP curriculum, with a contrasting profile of student performance resulting from the use of non-CMP curricula. Accordingly, the project was designed to answer three research questions:

1. What are the similarities and differences between the intended treatment of algebra in the CMP curriculum and in the non-CMP curricula?

2. What are key features of the CMP and non-CMP experience for students and teachers, and how might these features explain performance differences between CMP and non-CMP students?
3. What are the similarities and differences in performance between CMP students and a comparable group of non-CMP students on tasks measuring a broad spectrum of mathematical thinking and reasoning skills, with a focus on algebra?

In this article, our focus is on the third research question, even though we will highlight the differences between the CMP and non-CMP curricula, as well as the differences and similarities between CMP and non-CMP classroom instruction.

1.3. Focus on algebra

The main focus of the LieCal Project is to compare the effects of the CMP curriculum to the effects of non-CMP middle-school mathematics curricula on *students' learning of algebra*. We chose to compare the effects of the algebra strands of these curricula because of the importance of algebra in school mathematics. Middle school algebra lays the foundation for the acquisition of tools for representing and analyzing quantitative relationships, for solving problems, and for stating and proving generalizations (Bednarz, Kieran, & Lee, 1996; Cai & Knuth, 2011; Carpenter, Franke, & Levi, 2003; Cai and Knuth, 2011; Kaput, 1999; Mathematical Sciences Education Board, 1998; RAND Mathematics Study Panel, 2003). Algebra readiness has been characterized as the most important "gatekeeper" in school mathematics (Pelavin & Kane, 1990). Given its gatekeeper role as well as growing concerns about students' inadequate preparation in algebra in the United States, algebra curricula and instruction have become focal points of mathematics education research (Carpenter et al., 2003; Katz, 2007; National Research Council [NRC], 2004).

In research on algebra learning, more information is needed about the interplay between the acquisition of procedural knowledge and the acquisition of algebraic concepts (Kieran, 1997; National Academy of Education, 1999). By focusing on the algebra strand in the middle school we can examine explicit connections between the acquisition of algebraic concepts and the manner in which algebra is taught and learned. In addition, because algebra is both highly conceptual and highly procedural, this strand provides an interesting context within which to examine students' acquisition of both basic and higher-order thinking skills.

As we indicated above, in a *Standards-based* curriculum like CMP, the focus is on conceptual understanding and problem solving rather than on procedural knowledge.[1] Students are expected to learn algorithms and master basic skills as they engage in explorations of worthwhile problems. However many people, parents and teachers alike, worry that the development of students' higher-order thinking skills comes at the expense of fluency in computational procedures and symbolic manipulations. In summary, it is important to investigate not only how students develop higher-order thinking and basic skills, but also whether the use of the CMP curriculum comes at the expense of the development of computational and symbolic fluency.

In the current educational and political environment, there is an urgent need to understand the role that curriculum plays in students' learning of mathematics in general and in the acquisition of algebraic concepts in particular. Because *Standards*-based curricula like CMP not only look very different from commercially developed traditional mathematics curricula, but also claim to have different learning goals, they are well suited for examining the impact of curriculum on the development of students' algebraic thinking. By situating our examination of learning in a curricular context, we are able to investigate the role that curriculum plays in students' mathematics learning in general and in their acquisition of algebraic concepts in particular (NCTM, 1989; NRC, 2004; RAND Mathematics Study Panel, 2003; Senk & Thompson, 2003; Usiskin, 1999).

1.4. Classroom variables in examining curricular effect

Since the effectiveness of curricula depends critically on how well teachers implement them, studies of the effectiveness of *Standards*-based curricula must examine how teachers use the curricula (Kilpatrick, 2003; NRC, 2004; Wilson & Floden, 2001). The data gathered must be analyzed in appropriate ways to control for variations in classroom instruction and the learning environment. A study by Schoen, Cebulla, Finn, & Fi (2003) examined the relatedness of certain aspects of instructional practices to student achievement in high school classrooms in which a high school *Standards*-based curriculum was used. In particular, they used regression techniques to identify the teachers' background characteristics, behaviors, and concerns that are associated with growth in student achievement. They found that the percentages of class time spent on teacher presentation and on whole class discussion were each negatively associated with student achievement. However, the completion of a teachers' summer workshop on the use of the curriculum, the implementation of cognitively demanding tasks for students, and the adherence to reform principles during instruction were all significantly and positively associated with student achievement. A more recent study showed that coupling NSF-funded curricula with a *Standards*-based learning environment was associated with a significant positive impact on students' achievement (Tarr et al., 2008).

These studies confirm that in order to determine the effects of curriculum on learning, it is essential to examine the classroom experiences of the teachers and students who are using the different curricula. In another article, we analyzed the instructional tasks implemented in both CMP and non-CMP classrooms, and found that the tasks were more than three times

[1] In the revised version of the Connected Mathematics Program (called CMP 2), there is an increased emphasis on the procedural aspects of algebra.

as likely to be solved using multiple solution strategies in CMP classrooms as in non-CMP classrooms (Cai, Wang, Moyer, Nie, & Wang, submitted). In addition, we found that CMP teachers were more than three times as likely to implement high-level instructional tasks than non-CMP teachers and that the cognitive demand of the instructional tasks implemented in classrooms was a significant predictor of students' achievement gains over the three middle school years, regardless of the curriculum type.

In this article, we take features of classroom instruction into consideration when we examine the impact of curricula on students' learning of algebra. In particular, we examine the extent to which CMP and non-CMP teachers emphasize conceptual and procedural understanding in the classroom. As was reported by Moyer and his colleagues in the third article in this issue, CMP teachers placed more emphasis on conceptual understanding than non-CMP teachers. On the other hand, non-CMP teachers placed more emphasis on procedural understanding than CMP teachers. In this article, we particularly examine students' achievement gains across the three middle school years while controlling for the conceptual and procedural emphases in classroom instruction.

1.5. Equality

Success in algebra and geometry has been shown to help narrow the disparity between minority and non-minority participation in post-secondary opportunities (Loveless, 2008). Research shows that completion of an Algebra II course correlates significantly with success in college and with earnings from employment. The National Mathematics Advisory Panel (2008) found that students who complete Algebra II are more than twice as likely to graduate from college as students with less mathematical preparation. Furthermore, the African-American and Hispanic students who complete Algebra II cut the gap between their college graduation rate and that of the general student population in half. However, success in high school algebra is dependent upon mathematics experiences in the middle grades. In fact, middle school is a critical turning point for students' development of algebraic thinking (College Board, 2000). To what extent does the use of a *Standards*-based curriculum such as CMP improve the mathematics achievement for all students and help close achievement gaps between minority and non-minority students? This is an under-investigated question (Lubienski & Gutiérrez, 2008; Schoenfeld, 2002).

2. Differences between CMP and non-CMP curricula

Before we present evidence of the impact of curriculum reform on students' learning, we highlight some differences between the CMP curriculum and the non-CMP curricula. In particular, when we conducted detailed analyses of CMP and one of the non-CMP curricula[2], we found significant differences between them (Cai, Nie, & Moyer, 2010; Nie, Cai, & Moyer, 2009). Overall, our research revealed that the CMP curriculum takes a functional approach to the teaching of algebra, and the non-CMP curriculum takes a structural approach. The functional approach emphasizes the important ideas of change and variation in situations and contexts. It also emphasizes the representation of relationships between variables situated contextually. The structural approach, on the other hand, avoids contextual problems in order to concentrate on developing the abilities to generalize, work abstractly with symbols, and follow procedures in a systematic way (Cai et al., 2010). In this section, we highlight specific differences in the ways that the CMP curriculum and the non-CMP curriculum (1) define variables, (2) define equation solving, (3) introduce equation solving, and (4) use mathematical problems to develop algebraic thinking. We focus on these four aspects in this article because they are fundamental to algebra learning.

2.1. Defining variables

The learning goals of the CMP curriculum characterize variables as quantities used to represent relationships. In contrast, the learning goals in the non-CMP curriculum characterize variables as placeholders or unknowns. The CMP curriculum does not formally define variable until 7th grade. However, CMP's definition of variable as a quantity rather than a symbol makes it convenient to use variables informally to describe relationships long before formally introducing the concept of variable in 7th grade. Once CMP defines variables as quantities that change or vary, it uses them to represent relationships. The non-CMP curriculum, formally define a variable in 6th grade as a symbol (or letter) used to represent a number. It treats variables predominantly as placeholders and uses them mostly to represent unknowns in expressions and equations.

2.2. Defining equations

In CMP, the functional approach to equation is a natural extension of its development of the concept of variable as a changeable quantity used to represent relationships. At first, CMP expresses relationships between variables with graphs and tables of real-world quantities rather than with algebraic equations. Later, when CMP introduces equations, the emphasis is on using them to describe real-world situations. Rather than seeing equations simply as objects to manipulate,

[2] In this section, we only discuss the differences between CMP and one of the non-CMP curricula, Glencoe *Mathematics: Concepts and Applications* (Bailey et al., 2006a,b,c). Although there are differences among the non-CMP curricula, the differences between the CMP curriculum and each of the non-CMP curricula are similar.

Table 1
An example of equation solving in CMP (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2002b, p. 55).

The unlimited store allows any customer who buys merchandise costing over $30 to pay on the installment plan. The customer pays $30 down and then pays $15 a month until the item is paid for. Suppose you buy a $195 CD-ROM drive from the unlimited store on an installment plan, How many months will it take you to pay for the drive? Describe how you found your answer.

Thinking	Symbol manipulation
"I want to buy a CD-ROM drive that costs $195. To pay for the drive on the installment plan, I must pay $30 down and $15 a month."	$195 = 30 + 15N$
"After I pay the $30 down payment, I can subtract this from the cost. To keep the sides of the equation equal, I must subtract 30 from both sides	$195 - 30 = 30 - 30 + 15N$
"I now owe $165, which I will pay in monthly installments of $15″	$165 = 15N$
"I need to separate $165 into payments of $15. This means I need to divide it by 15. To keep the sides of the equation equal, I must divide both sides by 15″	$\frac{165}{15} = \frac{15N}{15}$
"There are 11 groups of $15 in $165, so it will take 11 months"	$11 = N$

The unlimited store allows any customer who buys merchandise costing over $30 to pay on the installment plan. The customer pays $30 down and then pays $15 a month until the item is paid for. Suppose you buy a $195 CD-ROM drive from the unlimited store on an installment plan, How many months will it take you to pay for the drive? Describe how you found your answer.

students learn that equations often describe relationships between varying quantities that arise from meaningful, contextualized situations (Bednarz, Kieran, & Lee, 1996). In fact, in CMP equations are formally defined as rules that are expressed with mathematical symbols, and that are often used for describing the relationship between two variables.

In the non-CMP curriculum, the definition of variable as a symbol develops naturally into two iconic hallmarks of a structural focus: the use of decontextualized (or "naked") equations and an emphasis on procedures for solving them. For example, immediately after defining an equation as "...a sentence that contains an equals sign,=," the non-CMP curriculum provides examples like $2 + x = 9$, $4 = k - 6$, and $5 - m = 4$ (Bailey et al., 2006a, p. 34). Students are told that the way to solve an equation is to replace the variable with a value that results in a true sentence.

2.3. Introducing equation solving

In the CMP curriculum, equation solving is introduced within the context of discussing linear relationships. The initial treatment of equation solving does not involve symbolic manipulation as found in most conventional curricula. Instead, the CMP curriculum introduces students to linear equation solving by making visual sense of what it means to find a solution using a graph. Its premise is that a linear equation in one variable is, in essence, a specific instance of a corresponding linear relationship (equation) in two variables. At first, equation solving relies heavily on the context within which the equation itself is situated, and on the use of a graphing calculator.

After CMP introduces equation solving graphically, the symbolic method of solving linear equations is eventually broached. It is introduced within a single contextualized example, where each of the steps in the equation solving process is accompanied by a narrative that demonstrates the connection between what is happening in the procedure and in the real-life situation. In this way, CMP justifies the equation-solving manipulations through contextual sense making of the symbolic method. That is, CMP uses real-life contexts to help students understand the meaning of each step of the symbolic method of equation solving, including why inverse operations are used, as shown in Table 1.

In the non-CMP curriculum, contextual sense making is not used to justify the equation-solving steps as it is in the CMP curriculum. Rather, the non-CMP curriculum first introduces equation solving as the process of finding a number to make an equation a true statement. Specifically, *solving* an equation is described as replacing a variable with a value (called the *solution*) that makes the sentence true. The process of equation solving is introduced in the non-CMP curriculum symbolically by using the additive property of equality (equality is maintained if the same quantity is added to or subtracted from both sides of an equation) and the multiplicative property of equality (equality is maintained if the same non-zero quantity is multiplied by or divided into both sides of an equation).

In the 6th grade, the Glencoe curriculum (Bailey et al., 2006a) formally introduces equation solving with inverse operations by way of an activity that uses a cup to stand for an unknown. The appropriate number of cups and counters used as manipulatives in the activity are initially positioned to exactly represent the equation's symbols. They are then used to illustrate each step of the symbolic manipulations (see Fig. 1).

Using manipulatives as described above is referred to as "Method 1″ and is typically shown adjacent to an example illustrating the corresponding solution using the strictly symbolic "Method 2." In this way, the non-CMP curriculum illustrates how each manipulative step is comparable to a symbolic step in a solution based on the algebraic properties of equality, which is shown through vertical work. Fig. 2 is an example of Method 2, showing how to solve a one-step equation.

2.4. Using mathematical problems

The extent of the differences between the CMP and non-CMP curricula can also be highlighted through an analysis of mathematical problems. Using a scheme developed by Stein & Lane (1996), we classified the mathematical tasks in the CMP

Table 2
Percentages of various tasks in CMP and non-CMP curricula.

	Memorization	Procedures without connections	Procedures with connections	Doing mathematics
CMP ($n = 920$)	0.43	27.93	61.52	10.11
Non-CMP ($n = 2391$)	4.60	74.57	18.24	2.59

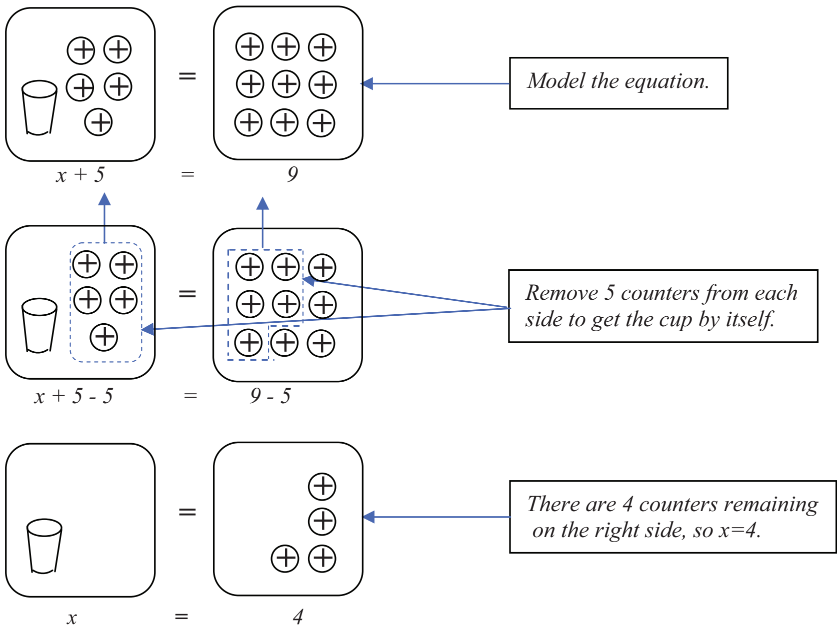

Fig. 1. The non-CMP curriculum (Bailey et al., 2006a) uses inverse operations to introduce equation solving.

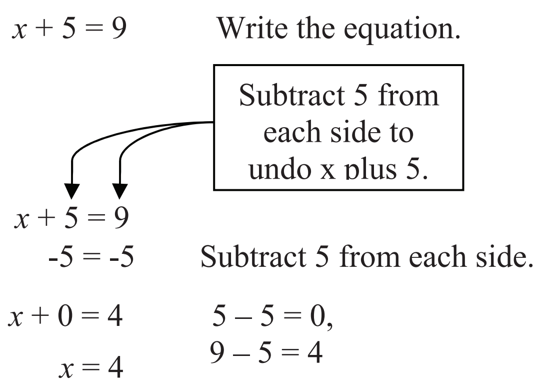

Fig. 2. The non-CMP (Bailey et al., 2006a) example of using symbolic representation to solve an equation.

curriculum and the non-CMP curriculum (Bailey et al., 2006a,b,c) into four increasingly demanding categories of cognition: memorization, procedures without connections, procedures with connections, and doing mathematics. As Table 2 shows, significantly more tasks in the CMP curriculum than in the non-CMP curriculum are higher-level tasks (procedures with connections and doing mathematics) ($\chi^2(3, N = 3311) = 759.52$, $p < .001$).

We further analyzed the problems in the CMP and non-CMP curricula that involve linear equations by classifying them into three categories:

1. One equation with one variable (1equ1va) – e.g., $2x + 3 = 5$.
2. One equation with two variables (1equ2va) – e.g., $y = 6x + 7$.
3. Two equations with two variables (2equ2va) – e.g., the system of equations $y = 2x + 1$ and $y = 8x + 9$.

Fig. 3 shows the percentage distribution of the problems involving linear equations in the two curricula. These two distributions are significantly different (χ^2 (2, $N = 2741) = 1262.0$, $p < .001$). The CMP curriculum includes a significantly

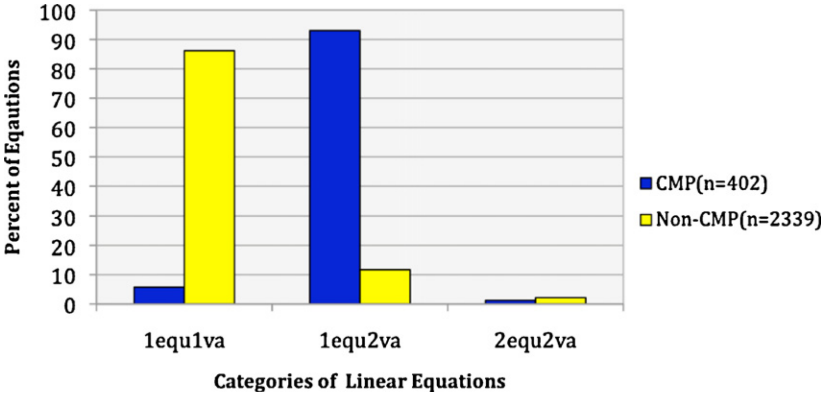

Fig. 3. Percentage distribution of problems involving linear equations in the CMP and non-CMP curricula.

greater percentage of "one equation with two variables" problems than the non-CMP curriculum ($z = 35.49$, $p < .001$). Also, the non-CMP curriculum includes a significantly greater percentage of "one equation with one variable" problems than the CMP curriculum ($z = 34.145$, $p < .001$). These results resonate with the findings that we reported above, namely that the CMP curriculum emphasizes an understanding of the relationships between the variables of equations, rather than an acquisition of the skills needed to solve them. In fact, of the 402 equation-related problems in the CMP curriculum, only 33 of them (about 8% of the linear equation solving problems) involve decontexualized symbolic manipulation of equation solving. However, the non-CMP curriculum includes 1550 problems involving decontexualized symbolic manipulation of equations (nearly 70% of the linear equation solving problems in the curriculum).

3. Methods

3.1. Sample

The LieCal project was conducted in 14 middle schools of an urban school district serving a diverse student population. When the project began, 27 of the 51 middle schools in the district had adopted the CMP curriculum, and the remaining 24 had adopted more traditional curricula. Seven schools were randomly selected from the 27 schools that had adopted the CMP curriculum. After the seven CMP schools were selected, seven non-CMP schools were chosen based on comparable demographics. In 6th grade, 695 CMP students in 25 classes and 589 non-CMP students in 22 classes participated in the study. We followed these 1284 students as they progressed from grades 6 to 8. Approximately 85% of the participants were minority students: 64% African American, 16% Hispanic, 4% Asian, and 1% Native American. The remaining 15% of the participants are Caucasians. Male and female students were almost evenly distributed.

3.2. Assessing students' learning

Learning algebra should involve much more than simply doing computations and solving equations. It should also provide students with a deep understanding of fundamental algebraic concepts, the connections between them, and the ability to use algebra to solve problems. The heart of measuring mathematical achievement is the set of tasks used to assess it (Mislevy, 1995; NRC, 2001). It is desirable to use various types of assessment tasks, thereby measuring different facets of algebraic thinking. Two important aspects of algebraic learning are conceptual understanding and problem solving, and symbol manipulation skills.

Table 3 summarizes our data collection. We used the state test scores in mathematics and reading as measures of prior achievement. We used the LieCal-developed multiple-choice and open-ended assessment tests as dependent measures of procedural knowledge and conceptual understanding in algebra, respectively. In this article, we report only the results from the two LieCal-developed tests, which we administered four times, each over two consecutive days.

Table 3
Data source and time of data collection.

Data sources	Fall, 05	Spring, 06	Fall, 06	Spring, 07	Fall, 07	Spring, 08
State tests on both math and reading	All 6th graders		All 7th graders		All 8th graders	
LieCal-developed test (multiple-choice items assessing procedural knowledge and basic skills in pre-algebra and algebra)	6th graders students (32 items)	6th graders students (32 items)		7th graders students (32 items)		8th graders students (32 items)
LieCal-developed test (open-ended tasks assessing conceptual understanding and higher order thinking skills in pre-algebra and algebra)	6th graders students (6 items)	6th graders students (5 items)		7th graders students (5 items)		8th graders students (5 items)

Table 4
Reliability coefficients for project-developed assessment forms.

	Multiple-choice	Open-ended form A	Open-ended form B	Open-ended form C
Fall 2005	.801	.647	.647	.647
Spring 2006	.836	.703	.719	.721
Spring 2007	.862	.764	.651	.710
Spring 2008	.876	.766	.743	.761

Note: Only one form of the open-ended assessment was administered in Fall 2005.

In our study, we used a combination of multiple-choice and open-ended assessment tasks to measure students' high-level thinking skills as well as students' procedural knowledge and routine problem-solving skills. We used multiple-choice items to assess whether students had learned the basic knowledge required to perform competently in introductory algebra. We decided to use multiple-choice questions because of their potential for broad content coverage and objective scoring, their highly reliable format, and their low cost of scoring. In addition to multiple-choice questions, we used open-ended assessment tasks. The open-ended tasks provided a better window than the multiple choice tasks into the thinking and reasoning processes involved in the students' algebra-related problem solving. The Appendix A shows sample items.

3.2.1. Multiple-choice items

While they were in grades 6–8, the LieCal participants were given four parallel versions of the multiple-choice test: F05 (baseline), Sp06, Sp07, Sp08. Each version comprised 32 questions that assessed five mathematics components: translation, integration, planning, execution (or computation), and equation solving. The items in the first four components are based on Mayer's (1987) model for analyzing cognitive components in solving word problems. Translation and integration involve the representing phase of problem solving, while planning and execution involve the searching phase of problem solving. In order to represent a problem, a student must be able to put the elements of a problem together into a coherent whole and translate them into an internal representation, such as an equation. In the searching phase of problem solving, the student must first plan the solution, and then find, and execute an adequate algorithm. In our multiple-choice test, we used six items to assess each of Mayer's four cognitive components, and 8 items to assess equation solving. Table 4 provides reliability coefficients (Cronbach's coefficients) for both multiple-choice and open-ended forms across assessment administrations.

3.2.2. Open-ended tasks

In addition to the baseline multiple-choice assessment administered in the fall of 2005, all the LieCal 6th-graders received a baseline open-ended assessment, which consisted of 6 tasks. These tasks, as well as the open-ended tasks used for later assessments, were adopted from various projects, including Balanced Assessment[3], the QUASAR Project (Lane et al., 1995), and a cross-national study (Cai, 2000). Since only a small number of open-ended tasks can be administered in a testing period, and since grading students' responses to such items is labor-intensive, we distributed the non-baseline open-ended tasks over three forms (five items in each form) and used a matrix sampling design to administer them. That is, starting in the spring of 2006, each third of the students was administered one of the three forms. In the springs of 2007 and 2008, the forms were rotated so that eventually each student received all three forms.

For this article, in addition to the results of the open-ended tasks, we report only the results from the translation, computation, and equation-solving components of the multiple-choice tasks. The open-ended tasks were designed to assess students' conceptual understanding and problem-solving skills. The translation component of the multiple-choice tasks assesses students' ability to represent problem situations. The items for the computation and equation-solving components of the multiple-choice tests assess students' procedural knowledge and symbol manipulation skills.

3.2.3. Scoring

The multiple-choice items were scored electronically, either right or wrong. The open-ended tasks were scored by middle school mathematics teachers, who were trained to score student responses using previously developed scoring rubrics. Two teachers scored each response. On average, perfect agreement between each pair of raters was nearly 80%, and agreement within one point difference out of 6 points (on average) was over 95% across tasks. Differences in scoring were arbitrated through discussion.

3.2.4. Linking items and scaled scores

From one testing administration to another, 10 of the 32 multiple-choice items were identical, while the other 22 items were new, but parallel. The 10 identical items comprised two items from each of the five components. They served as "linking items" in the analysis. In a similar way, at least two identical open-ended tasks served as linking items from one form to another and one testing administration to another. We used scaled scores to report and analyze the student achievement data. A scaled score

[3] We thank Alan Sheinker for allowing us to use a few Balanced Assessment items, published by CTB/McGraw-Hill. The development of Balanced Assessment was led by Sandra Wilcox, Michigan State University; Alan Schoenfeld, University of California, Berkeley; Hugh Burkhardt, Shell Centre, University of Nottingham, England; Jim Ridgway, University of Durham, England; and Phil Daro, University of California, Chancellor's Office.

is a generic term for a mathematically transformed student raw score on an assessment. Using scaled scores, rather than raw scores, made it possible to place assessment results on the same scale even if students responded to different tasks at different times. The two-parameter Item Response Theory (IRT) model was used to scale student assessment data on both multiple-choice tasks and open-ended tasks (Hambleton, Swaminathan, & Rogers, 1991; Lord, 1980).

3.3. Conceptual and procedural emphases as classroom variables

In this article, we use conceptual and procedural emphases as classroom variables to examine the impact of curriculum on students' learning. We estimated the levels of conceptual and procedural emphases in the CMP and non-CMP classrooms using data from 620 lesson observations of the LieCal teachers, which we conducted while the students were in grades 6, 7, and 8. The details surrounding the observations are documented in Moyer et al. (this issue). Since students changed their classrooms and teachers as they moved from grade 6 to grade 7 and from grade 7 to grade 8, all students in the same classroom at each grade were measured and had the same value but each student could have a different value each year for three years. One component of the observation data is a set of 21 questions using a 5-point Likert scale that were used to rate the nature of instruction for each lesson. Of the 21 questions, four of them are designed to assess the extent to which a teacher's lesson has a conceptual emphasis. Another four of the questions are designed to determine the extent to which a teacher's lesson has a procedural emphasis. Factor analysis of the LieCal observation data confirmed that the four procedural-emphasis questions loaded on a single factor, as did the four conceptual-emphasis questions.

There was a significant difference across grade levels among the levels of conceptual emphasis in the CMP and non-CMP instruction ($F(3, 575) = 53.43$, $p < .001$). The overall mean (grades 6–8) of the summated ratings of conceptual emphasis in CMP classrooms was 13.41 with a standard deviation of 3.76, while the overall mean of the summated ratings of conceptual emphasis in non-CMP classrooms was 10.06 with a standard deviation of 2.55.

On the other hand, non-CMP lessons had significantly more emphasis on the procedural aspects of learning than did the CMP lessons ($F (3, 575) = 37.77$, $p < .001$). Also, the overall mean (grades 6–8) of the summated ratings of procedural emphasis in non-CMP classrooms was 14.49 with a standard deviation of 3.44, while the overall mean of the summated ratings of procedural emphasis in CMP classrooms was 11.61 with a standard deviation of 3.18.

3.4. Analysis of achievement data

We analyzed the longitudinal data using two complementary statistical approaches. First, we analyzed the longitudinal effect of curriculum on student achievement using a repeated measures analysis of variance (ANOVA). Second, because the data collected in the project is hierarchical in nature, we also used multilevel statistical models to capture student achievement changes over time and to analyze the longitudinal effects of the CMP and non-CMP curricula on students' learning (Raudenbush & Bryk, 2002). In particular, we used growth-curve modeling to examine the longitudinal effect of curriculum while taking into account both student-level variables (e.g., gender and ethnicity) and classroom variables (conceptual and procedural emphases).

For the growth-curve modeling, initially we performed three-level analyses by nesting students within teachers and teachers within schools. A complicating factor was that the vast majority of students had different teachers in grades 6, 7, and 8. We handled this complication by averaging the levels of conceptual emphasis (and also procedural emphasis) across the three years, nesting students within level of conceptual emphasis for one analysis and within level of procedural emphasis for the other. This was justified because the levels at which the instruction emphasized conceptual understanding or procedural understanding were the only instructional factors that we were interested in using as independent variables at the teacher/classroom level. It turned out, however, that the intraclass correlation coefficient (ICC), or percentage of variance between teachers/classrooms, was very small for all outcome variables (Kreft & de Leeuw, 1998). Table 5 gives the details.

Based on this result, we changed the growth curve model to a two-level model with the mean[4] of conceptual emphasis or procedural emphasis across three years as a classroom variable together with student ethnicity and curriculum type nested in schools.

The two-level simple conditional growth curve model is as follows[5]:

[4] All students in the same classroom were assigned the same value for the conceptual emphasis variable and the same value for the procedural emphasis variable. Since students changed their classrooms and teachers as they moved from grade 6 to grade 7 and from grade 7 to grade 8, many students were assigned a different value for the conceptual emphasis variable at each grade and a different value for the procedural emphasis at each grade. When the level of a classroom variable changes over time, it is common to use the average to represent the classroom variable (e.g., Domitrovich, Gest, Gill, Bierman, Welsh, & Jones, 2009; Reardon & Galindo, 2009). We are fully aware, however, that the use of average cannot distinguish students who, for example, were in classrooms with two high conceptual emphasis scores and one low conceptual emphasis score from students who were in classrooms with three medium scores. That is, students whose 6-8 classrooms were rated as High–High–Low would have a similar conceptual emphasis mean score as the students whose classrooms were rated as Medium–Medium–Medium. But these students would have had very different classroom experiences. Luckily, we found that only a small proportion (less than 10%) of the students who attended classes rated at the two extremes (high and low). Nonetheless, we used another procedure to run the growth curve modeling, which treated the classroom variable (conceptual emphasis or procedural emphasis) as time-variant in the Level One model ($Y_{ijk} = \pi_{0jk} + \pi_{1jk}t_{ijk} + \pi_{2jk}Concept_{ijk} + \pi_{3jk}Procedure_{ijk} + e_{ijk}$). When we did so, we found the same results as presented in this article, which used the mean of the of the classroom variables in the model.

[5] Unconditional models were always run before conditional models, but these models are not provided here because they are simply models without any of the second-level predictors that can be inferred from the conditional models.

Table 5
Intra-class correlation coefficients for outcome measures.

Outcome measure	ICC for the intercept	ICC for the slope
Open-ended tasks	4.47%	8.53%
Translation component	4.17%	5.09%
Computation component	3.52%	3.73%
Equation solving component	9.16%	0.60%

Level 1:

$$Y_{ti} = \pi_{0i} + \pi_{1i}t_{ti} + e_{ti}$$

where Y_{ti}, is the outcome variable on each of the dependent measures (e.g., score on open-ended tasks) at time t for student i, π_{0i}, is the initial status of student i on each of the dependent variables (e.g., predicted score on open-ended tasks) for that student at the beginning of the study; π_{1i}, is the constant annual growth rate for student i during the three years; t_{ti}, takes on a value of 0 in the beginning (Fall 2005), a value of 1 at the end of the first year (Spring 2006), a value of 2 at the end of the second year (Spring 2007), and a value of 3 at the end of the third year (Spring 2008).

Level 2: Each of the independent variables is used to predict the coefficients in the Level 1 Model. The following is the simple conditional model using *CMP* as the predictor during the first step of our HLM analysis. A similar model for each of the independent variables was used at the first step.

$$\pi_{0i} = \beta_{00} + \beta_{01}(CMP)_i + r_{0i}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}(CMP)_i + r_{1i}$$

Since "CMP" is an indicator variable (i.e. taking on the values 0 or 1), the corresponding regression coefficients can be interpreted as treatment effects. That is, β_{01} is the difference in the initial status (i.e., the extent to which an average non-CMP student starts ahead/behind an average CMP student) on the outcome variable. β_{11} represents the gap in the annual growth rates (i.e., the difference between average CMP and average non-CMP students in subsequent growth rates).

As implied above, a two-step strategy (Compton, 2000) for the conditional models was used. First, simple conditional models were run to examine each independent variable individually. Second, the variables significant ($p < .05$) at the first step were examined simultaneously (complete conditional model). The following is a complete conditional model at Level 2 if all variables in simple conditional models were statistically significant:

$$\pi_{0i} = \beta_{00} + \beta_{01}(CMP)_i + \beta_{02}(Male)_i + \beta_{03}(A\,fricanAmerican)_i + \beta_{04}(His\,panic)_i +$$
$$\beta_{05}(Caucasian)_i + \beta_{06}(Asian)_i + \beta_{07}(Conce\,pt)_i + \beta_{08}(Procedure)_i + r_{0i}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}(CMP)_i + \beta_{12}(Male)_i + \beta_{13}(A\,fricanAmerican)_i + \beta_{14}(His\,panic)_i +$$
$$\beta_{15}(Caucasian)_i + \beta_{16}(Asian)_i + \beta_{17}(Conce\,pt)_i + \beta_{08}(Procedure)_i + r_{1i}$$

The variable *Concept* (conceptual emphasis) was measured on an ordinal scale that has five levels (1 = low level, 3 = median level, 5 = high level). This is the same for procedural emphasis. All other independent (predictor) variables are dichotomous (0 = no and 1 = yes) so that each b_{0i} and b_{1i} coefficient (i = 2 . . . 7) represents the difference between two groups on initial status and growth rate, respectively. Grand-mean centering was used for concept and procedure, so the intercept represents the adjusted mean outcome score (e.g., open-ended tasks) for students whose concept and procedure were at the mean of the whole sample in the fall of 2005. No centering was used for dichotomous variables, so the intercept represents the adjusted mean of students who were coded "zero" on these dichotomous variables in the fall of 2005. The slope represents the estimated mean annual growth for the students who were coded "zero" on the dichotomous independent variables and whose concept and procedure were at the mean of the whole sample.

Since we were interested in whether or not the CMP curriculum had a differential impact on students of different ethnic groups, we also did an additional analysis for CMP students only by examining potential differences between students of different ethnic groups. Magnitude of effect, or proportion of variance explained by the complete model, was calculated by 1 minus the ratio between the estimated variance of the complete conditional model and that of the unconditional model.

4. Results

We first present descriptive statistics, then the findings from the longitudinal analysis using both repeated measures ANOVA and growth curve modeling. For the growth curve modeling, we did the analysis four times, once for each of the four dependent measures: (1) the open-ended tasks, (2) the translation component, (3) the computation component, and (4) the equation-solving component of the multiple-choice tasks. The open-ended tasks measure conceptual understanding and problem solving. The translation component of the multiple-choice tasks measures the students' ability to represent problem situations, while the computation and equation solving components measure symbol manipulation skills.

Table 6
Descriptive statistics for the four dependent measures in Fall '05, Spring '06, Spring '07, and Spring '08.

	Fall'05			Sp'06			Sp'07			Sp'08		
	Mean*	SD	*n*	Mean	SD	*n*	Mean	SD	*n*	Mean	SD	*n*
Open-ended												
CMP	459	84	629	471	96	606	518	92	621	549	101	723
Non-CMP	468	88	543	485	96	561	516	95	541	542	105	517
Translation												
CMP	436	87	623	485	94	615	519	95	614	540	96	712
Non-CMP	453	89	525	495	91	562	532	96	529	538	96	521
Computation												
CMP	466	89	623	482	99	615	511	102	614	518	98	712
Non-CMP	481	89	525	486	96	562	531	104	529	534	105	521
Equation-solving												
CMP	450	81	623	457	77	615	493	93	614	527	104	712
Non-CMP	496	94	525	501	97	562	539	101	529	563	101	521

*The score range is 200–800.

4.1. Descriptive statistics

Table 6 provides descriptive statistics for the four dependent measures in the Fall 2005, Spring 2006, Spring 2007, and Spring 2008 assessments. For each measure, the mean scaled scores increased from Fall 2005 to Spring 2008, for both CMP and non-CMP students.

4.1.1. Gains

We determined and ranked each student's gain score on each of the outcome measures from 6th to 8th grade. Table 7 shows the mean gains and the percentage of CMP and non-CMP students who had positive gains on each of the learning outcome measures. Table 7 also shows the percentages of CMP and non-CMP students whose gains ranked in the top or bottom 25% of all gains on each of the outcome measures over the three years.

For the open-ended tasks, CMP students had significantly higher mean gains than non-CMP students ($t = 2.20$, $p < .05$). In addition, a significantly larger percentage of the CMP students than non-CMP students had positive gains ($z = 2.71$, $p < .01$). The percentage of CMP students (14%) whose gain scores on the open-ended tasks rank in the top 25% was higher than the non-CMP students (11%) ($z = 2.57$, $p < .01$). On the other hand, the percentage of non-CMP students (14%) whose gain scores on the open-ended tasks rank in the bottom 25% was higher than the CMP students (11%) ($z = 2.16$, $p < .05$).

For the translation component of the multiple-choice tasks, CMP students had significantly higher mean gains than non-CMP students ($t = 2.57$, $p < .05$). However, there was no significant difference between the percentages of the CMP students and non-CMP students who had positive gains. The percentage of CMP students (15%) whose gain scores on the translation component rank in the top 25% was significantly higher than the non-CMP students (10%) ($z = 3.51$, $p < .01$). However, there was no difference between the percentage of the CMP students and non-CMP students whose Translation gain scores ranked in the top 25%.

Table 7
Mean gains and percentages of students with various gains from 6th grade to 8th grade by curriculum type.

	CMP	Non-CMP
Open-ended tasks		
Mean gains*	90.25	70.03
% of students with positive gains*	89%	83%
% of students in the top 25% of gains*	14%	11%
% of students in the bottom 25% of gains*	11%	14%
Translation		
Mean gains*	108.39	88.83
% of students with positive gains	85%	80%
% of students in the top 25% of gains*	15%	10%
% of students in the bottom 25% of gains	12%	13%
Computation		
Mean gains	49.64	62.59
% of students with positive gains*	60%	78%
% of students in the top 25% of gains	12%	13%
% of students in the bottom 25% of gains*	15%	10%
Equation solving		
Mean gains	67.04	76.16
% of students with positive gains	70%	73%
% of students in the top 25% of gains	12%	13%
% of students in the bottom 25% of gains	13%	12%

* $p < 0.05$.

For the computation component of the multiple-choice tasks, the mean gains of CMP (49.64) and non-CMP students (62.59) were not significantly different. However, a significantly larger percentage of the non-CMP students than CMP students had positive gains ($z = -2.94$, $p < .01$). There was no difference between the percentages of the CMP students and non-CMP students whose Computation gains rank in the top 25%. However, the percentage of CMP students (15%) whose gain scores on Computation rank in the bottom 25% was significantly higher than the non-CMP students (10%) whose gain scores on Computation rank in the bottom 25%, ($z = 4.22$, $p < .001$).

For the equation-solving component of the multiple-choice tasks, there was no significant difference between the mean gains of CMP and non-CMP students from 6th grade to the 8th grade, nor was there a significant difference between the percentages of the CMP and non-CMP students who had positive gains over the three years. In addition, there was no significant difference between the percentages of the CMP students and non-CMP students whose Equation Solving gains rank in the top 25%, nor was there a significant difference between the percentages of CMP and non-CMP students whose gains rank in the bottom 25%.

4.2. Repeated measures ANOVA

As mentioned earlier, we used a repeated measures ANOVA to analyze the longitudinal effect of curriculum on the four dependent measures of student achievement. Table 8 shows the *F*-values for the main effect of time, the main effect of curriculum, and the time and curriculum interaction on the four dependent measures.

For the repeated measures ANOVA, we could only use data from the cohort of students who took all four assessments (Fall 2005, Spring 2006, Spring 2007, and Spring 2008). We conducted Chi-square analyses to examine whether the students who took all four project-developed assessments had ethnicity characteristics that were similar to those in the initial sample (i.e., to the 6th grade students who started in the project). The results indicated that the ethnic distribution of the reduced sample used in the analyses was not statistically different from that of the original sample. On the open-ended tasks, the Chi-square analysis yielded $\chi^2 = 1.93$ ($p = 0.75$) for the CMP students and $\chi^2 = 2.95$ ($p = 0.57$) for the non-CMP students. On the multiple-choice tasks (which include the translation, computation, and equation solving components), the Chi-squares were $\chi^2 = 1.36$ ($p = 0.85$) and $\chi^2 = 1.55$ ($p = 0.82$) for the CMP and non-CMP students, respectively. These results imply that the test-taking attrition was proportionately equal across ethnic groups. Therefore, even though the data we used in the analyses were from only a subset of the LieCal students, the subset of the students has characteristics similar to the entire cohort.

4.2.1. Open-ended tasks

Table 8 shows a significant main effect due to time on the open-ended tasks. From Fall 2005 (6th grade) to Spring 2008 (8th grade), both CMP and non-CMP students showed significant growth on the open-ended tasks ($F(3, 602) = 305.326$, $p < .001$). There was no main effect due to curriculum on the open-ended tasks. This suggests that overall CMP and non-CMP students performed equally well on the open-ended tasks. However, there is a significant interaction between time and curriculum. Coupled with an examination of the mean gains on the open-ended tasks, as shown in Table 7, this significant interaction suggests that the annual growth rates of the CMP students was significantly higher than that of the non-CMP students ($F(3, 602) = 3.341$, $p < .05$).

4.2.2. Translation tasks

From Fall 2005 (6th grade) to Spring 2008 (8th grade), both CMP and non-CMP students showed significant growth on the translation component of the multiple-choice tasks (F (3, 572) = 219.681, $p < .001$), which measures students' ability to represent situations. The repeated measures ANOVA did not reveal a main curriculum effect, which suggests that the CMP and non-CMP students performed equally on the translation tasks. However, as Table 8 shows, there was a significant time and curriculum interaction on translation ($F(3, 572) = 2.822$, $p < .05$). This result suggests that the CMP students show a significantly higher growth rate than the non-CMP students on representing-situations items. As in the open-ended tasks, the CMP students started lower than the non-CMP students in the fall of 2005, but by the spring of 2008, they performed better than the non-CMP students on the translation tasks.

4.2.3. Computation tasks

Although both CMP and non-CMP students showed significant growth on the computation tasks ($F(3, 572) = 100.435$, $p < .001$) from Fall 2005 to Spring 2008, there was no main curriculum effect and no significant interaction between time and

Table 8
F-values of repeated measures ANOVA on the dependent measures.

	Open-ended	Translation	Computation	Equation Solving
Time	305.326**	219.681**	100.435**	139.300**
Curriculum	0.018	0.455	1.949	58.789**
Time × curriculum	3.341*	2.822*	2.533	0.589

* $p < .05$.
** $p < .001$.

curriculum. This suggests that the CMP students did not perform differently than the non-CMP students on the computation tasks. It also suggests that the CMP and non-CMP students showed similar growth rates across the three years (grades 6–8).

4.2.4. Equation-solving tasks

Similar to the open-ended, translation, and computation tasks, both CMP and non-CMP students showed significant growth on the equation solving tasks ($F(3, 572) = 139.300$, $p < .001$) from Fall 2005 to Spring 2008. However, unlike on the other three types of tasks, there was a significant main effect of curriculum on equation solving. Coupled with an examination of the mean gains on the equation-solving tasks shown in Table 7, the significant main curriculum effect on Equation Solving suggests that non-CMP students performed better than CMP students on the equation solving tasks. As with the computation tasks, there was no significant interaction between time and curriculum. This suggests that CMP and non-CMP students had similar growth rates over the four testing administrations from Fall 2005 to Spring 2008.

4.3. G growth curve modeling

One of the advantages of using the growth curve modeling is to test the significance of slopes across time and variables predicting the change of slopes (Raudenbush, Bryk, & Congdon, 2005). Thus, we used the growth curve modeling to examine the effect of curricula, while controlling instructional and student variables on the four dependent measures. The four dependent measures, Open-Ended Tasks, Translation, Computation, and Equation Solving, were each analyzed separately. As we described above, a two-level growth curve model was used. The first level is an individual growth model (time was coded as 0, 1, 2, and 3 so that the slope represents annual growth), and the parameters in the Level 1 model are estimated by the outcome variables in the Level 2 model. Student characteristics (curriculum status, gender, and ethnicity), and instructional variables (conceptual and procedural emphases) were used as the predictors in the second level.

There were 2001 students who had at least one data point on each of the dependent measures. Since HLM allows unequal numbers of data points per student as well as unequal spacings of data points over time (Raudenbush & Bryk, 2002), all these 2001 students were included in the analyses using growth curve modeling. However, students whose data were missing for more than one of the 4 measures (*F* '05, Sp '06, Sp '07, Sp '08) of a particular dependent variable were excluded by the HLM software from the Level 1 analysis for that particular dependent variable (but were not necessarily excluded from the Level 1 analysis for the other dependent variables). This is because the growth curve model needs at least 3 measures. The end result was a Level 1 sample size of 1315 for the open-ended tasks and a sample size of 1345 for each of the three multiple-choice components.

As we have indicated, at Level 2 we used the dependent-variable/independent-variables variance-covariance matrix to predict the coefficients at Level 1. All the students who had at least one valid data point on any of the dependent and independent variable(s) were included in the Level 2 analysis. The sample size was 1740 when Open-Ended Tasks was the dependent variable in the Level 2 model, but the sample size was 1729 when each of translation, computation, and equation solving was the dependent variable.

4.3.1. Open-ended tasks

Similar to what we found in the repeated measures ANOVA, the analysis using growth curve modeling showed that the CMP students started a bit lower in their performance in the fall of 2005 than the non-CMP students, but not significantly so, as seen in the results from the simple conditional model shown in Table 9. Moreover, African American students scored lower

Table 9
Open-ended total scaled score as the dependent variable.

Estimated parameters	Simple conditional model				Complete conditional model			
	Coefficient	SE	t	p	Coefficient	SE	t	p
Intercept					454.94	3.97	114.64	<.001
CMP	−8.21	4.60	−1.78	.07				
Gender	4.08	4.59	0.88	.38				
African American	−41.36	4.68	−8.83	<.001	−19.82	4.59	−4.32	<.001
Hispanic	−3.86	5.39	−0.72	.47				
Caucasian	89.05	7.93	11.22	<.001	66.35	7.90	8.40	<.001
Asian	22.22	11.05	2.01	.04	11.44	11.53	0.99	.32
Slope					19.39	2.64	7.36	<.001
CMP	4.61	1.65	2.79	.006	5.70	1.58	3.61	.001
Gender	−2.46	1.64	−1.51	.13				
African American	−3.13	1.67	−1.88	.06				
Hispanic	3.56	2.09	1.70	.09				
Caucasian	−1.37	2.45	−0.56	.58				
Asian	9.76	3.60	2.71	.007	9.59	3.65	2.63	.01
Concept	1.82	1.21	1.50	.13				
Procedure	−2.11	1.40	−1.50	.13				

($t = -8.83$, $p < .001$) while Caucasian ($t = 11.22$, $p < .001$) and Asian students ($t = 2.01$, $p = .04$) scored higher than students of other ethnicities in the fall of 2005. However, the simple conditional model also suggests that, over the three middle school years, the CMP students' scores increased significantly more than the non-CMP students' scores ($t = 2.79$, $p < .05$).

The simple conditional model also shows that there is no statistical difference between the growth rates of African Americans and non-African Americans, nor between Hispanics and non-Hispanics or Caucasians and non-Caucasians. Asian students experienced significantly higher annual growth rates than students of other ethnic groups ($t = 2.71$, $p = .007$). Finally, the simple conditional model shows that the level of conceptual emphasis or procedural emphasis in instruction did not have a statistically significant impact on the growth rate of students' scores.

The complete conditional model suggests that the estimated mean score for non-CMP female students who were non-African American, non-Caucasian, and non-Asian was 454.94 in the fall of 2005 and that the estimated mean annual growth rate for non-CMP and non-Asian students was 19.39. These estimated mean score and annual growth rate were significantly different from zero ($t = 114.64$, $p < .001$ and $t = 7.36$, $p < .001$, respectively). CMP students' growth rate was still significantly higher than non-CMP students when students' ethnicity was controlled ($t = 3.61$, $p < .01$). In particular, CMP students had an annual gain of 19.39 + 5.70 = 25.09 scale points whereas non-CMP students had an annual gain of 19.39. Asian students in the CMP curriculum had an annual growth rate of 19.39 + 5.70 + 9.59 = 34.68 scale points. The magnitude of effect of this complete model was 30%.

Additional analysis of students of different ethnic groups within the CMP curriculum (not shown in Table 9) revealed that Hispanic students in the CMP program benefited more than other students. That is, their growth rate was significantly higher than that of non-Hispanic students ($t = 2.07$, $p < .05$). The students in the CMP program did not differ significantly from that of CMP non-African American students. Neither were there differences between the growth rates of the CMP Asian and non-Asian students, nor between the growth rates of the CMP Caucasian and non-Caucasian students. These four results indicate that the CMP program did not have a negatively biased impact on the growth rates of any of the ethnic groups, but more notably not on the African American and Hispanic students' growth rates.

4.3.2. Translation tasks

The simple conditional model shown in Table 10 suggests that CMP student scored significantly lower than non-CMP students in the fall of 2005 ($t = -3.70$, $p < .001$). Similarly, African American students ($t = -2.87$, $p = .01$) and Hispanic students ($t = -2.11$, $p = .04$) scored significantly lower than non-African American students and non-Hispanic students. Caucasian students, however, scored significantly higher than non-Caucasian students in the fall of 2005 ($t = 7.56$, $p < .001$). CMP students had a significantly higher growth rate than non-CMP students ($t = 2.24$, $p < .05$). Furthermore, the model shows that the level of conceptual emphasis in instruction had a positive impact on the growth rate of students' performance ($t = 2.79$, $p < .05$). In fact, we found that with a unit increase in the level of conceptual emphasis, the students' growth rate will increase by 4.26 scaled-score points per year. Table 10 also shows, however, that the level of procedural emphasis in instruction did not have a statistically significant impact on the growth rate of students' performance. Further, the simple conditional model found no significant difference between the growth rates of male and female students.

The simple conditional model also shows that there was no significant difference between the performance on the translations tasks of students from any ethnic group and the rest of the students. As Table 10 shows, $t = 0.02$ ($p = .99$) for African American and non-African American students; $t = -0.74$ ($p = .46$) for Hispanic and non-Hispanic students; $t = 0.50$ ($p = .62$) for Caucasian and non-Caucasian students; and $t = 0.66$ ($p = .51$) for Asian and non-Asian students.

The complete conditional model suggests that the estimated mean score of non-CMP, non African American, Non-Hispanic, and non-Caucasian students was 447.11 in the fall of 2005 and the estimated mean annual growth rate of non-CMP

Table 10
Translation total scaled score as the dependent variable.

Estimated parameters	Simple conditional model				Complete conditional model			
	Coefficient	SE	*t*	*p*	Coefficient	SE	*t*	*p*
Intercept					447.11	10.72	41.69	<.001
CMP	−17.37	4.69	−3.70	<.001	−3.22	5.19	−0.62	.54
Gender	−4.69	4.71	−1.00	.32				
African American	−13.60	4.75	−2.87	.01	−1.11	6.77	−0.16	.87
Hispanic	−11.29	5.34	−2.11	.04	−9.21	7.32	−1.26	.21
Caucasian	56.51	7.48	7.56	<.001	47.80	8.41	5.68	<.001
Asian	−15.61	10.13	−1.54	.12				
Slope					27.70	4.06	6.82	<.001
CMP	4.82	2.15	2.24	.03	3.54	2.58	1.38	.17
Gender	−3.00	2.17	−1.39	.17				
African American	0.04	2.16	0.02	.99				
Hispanic	−1.89	2.58	−0.74	.46				
Caucasian	1.57	3.13	0.50	.62				
Asian	3.07	4.64	0.66	.51				
Concept	4.26	1.52	2.79	.01	2.15	1.59	1.35	.18
Procedure	−1.16	1.83	−0.64	.53				

students whose concept is at the mean of the whole sample was 27.70. The estimated mean score and annual growth rate were significantly different from zero ($t = 41.69$, $p < .001$ and $t = 6.82$, $p < .001$, respectively). The difference between CMP and non-CMP students' growth rate on translation tasks diminished after controlling the instructional variables. That is to say, when the teacher's conceptual and procedural emphases were at the same level, there was no difference between CMP and non-CMP students with respect to their growth on translation tasks. Similarly, the impact of the level of conceptual understanding on instruction also became insignificant when the students' curriculum status was controlled. The magnitude of effect of this complete model was 26%.

Additional analysis of students of different ethnic groups within the CMP curriculum (not shown in Table 10) revealed that the CMP curriculum did not have a biased/different impact on the growth rates of students of differing ethnicities.

4.3.3. Computation tasks

As Table 11 shows, the simple conditional model suggests that African American students ($t = -5.47$, $p < .001$) scored less whereas Caucasian ($t = 7.42$, $p < .001$) and Asian students ($t = 2.56$. $p = .01$) students scored higher than students of other ethnic groups in the fall of 2005. In addition, students in CMP classes had lower growth rates than students in non-CMP classes ($t = -1.95$, $p < .05$). There was no statistically significant difference between male and female students' growth rates ($t = 0.52$, $p = .60$). African American students had a significantly lower growth rate in comparison with students of other ethnic groups ($t = -4.75$, $p < .001$). Hispanic, Caucasian, and Asian students all had significantly higher growth rates in comparison to students outside their ethnic groups. (As Table 11 shows, the t-values and significance levels were $t = 2.17$ ($p < .05$) for Hispanic and non-Hispanic students; $t = 2.64$ ($p < .05$) for Caucasian and non-Caucasian students; and $t = 2.67$ ($p < .05$) for Asian and non-Asian students.) The amount of instructional emphasis placed on conceptual understanding did not have a statistically significant impact on the growth rate, nor did the amount of instructional emphasis placed on procedures.

The complete conditional model suggests that the estimated mean score for non-African American, non Caucasian, and non-Asian students was 462.87 in the fall of 2005 and the estimated mean annual growth rate for non-CMP, non-African American, non-Hispanic, non-Caucasian, and non-Asian students was 33.81. These estimated mean score and annual growth rate were significantly different from zero ($t = 108.27$, $p < .001$ and $t = 6.22$, $p < .001$, respectively). Students in CMP classes had significantly lower growth rates than students in non-CMP classes when students' ethnicity was controlled ($t = -2.38$, $p < .05$). African American students had a significantly lower growth rate in comparison with students of other ethnic groups when students' curriculum status was controlled ($t = -2.41$, $p < .05$). Non-African American students in CMP classes had an annual increase on scaled scores of $33.81 - 4.29 = 29.52$, whereas non-African American students in non-CMP classes had an annual increase in scaled scores of 33.81. African American students in CMP classes had an annual growth rate of $33.81 - 12.10 - 4.29 = 17.42$. In comparison, African American students in non-CMP classes had an annual increase on computation scaled scores of $33.81 - 12.10 = 21.71$. The magnitude of effect of this complete model was 12%.

Additional analysis of students of different ethnic groups within the CMP curriculum (not shown in Table 11) revealed that African American students in the CMP program benefited less than other students. Their growth rate was significantly lower than that of other students ($t = -3.09$, $p < .001$). Students of other ethnic groups (Hispanic, Caucasian, and Asian) in the CMP program were not found to differ significantly from others with respect to the growth rate, indicating that the CMP program did not have a negatively biased impact on the growth rate of any of the non-African American ethnic groups. Of particular interest is the fact that there was no indication that the CMP program had a negative impact on the Hispanic students' growth rate.

Table 11
Computation total scaled score as the dependent variable.

Estimated parameters	Simple conditional model				Complete conditional model			
	Coefficient	SE	t	p	Coefficient	SE	t	p
Intercept					462.87	4.28	108.27	<.001
CMP	−7.18	4.77	−1.51	.13				
Gender	−1.70	4.79	−0.36	.72				
African American	−26.31	4.81	−5.47	<.001	−4.57	5.26	−0.87	.39
Hispanic	−6.20	5.36	−1.16	.25				
Caucasian	59.20	7.98	7.42	<.001	45.66	8.82	5.18	<.001
Asian	26.60	10.38	2.56	.01	20.79	10.81	1.92	.05
Slope					33.81	5.43	6.22	<.001
CMP	−3.93	2.01	−1.95	.05	−4.29	1.80	−2.38	.02
Gender	1.05	2.03	0.52	.60				
African American	−9.53	2.01	−4.75	<.001	−12.10	5.02	−2.41	.02
Hispanic	5.28	2.44	2.17	.03	−4.58	5.19	−0.88	.38
Caucasian	7.15	2.71	2.64	.01	−1.63	5.43	−0.30	.76
Asian	12.72	4.76	2.67	.01	5.29	6.67	0.79	.43
Concept	1.17	1.57	0.74	.46				
Procedure	2.88	1.63	1.77	.08				

Table 12
Equation total scaled score solving as the dependent variable.

Estimated parameters	Simple conditional model				Complete conditional model			
	Coefficient	SE	t	p	Coefficient	SE	t	p
Intercept					501.62	9.68	51.82	<.001
CMP	−39.67	4.64	−8.55	<.001	−24.79	5.59	−4.43	<.001
Gender	−1.04	4.70	−0.22	.83				
African American	−27.85	4.75	−5.86	<.001	−13.83	4.07	−3.40	.001
Hispanic	1.16	5.29	0.22	.83				
Caucasian	56.58	8.08	7.00	<.001	35.10	8.17	4.30	<.001
Asian	1.51	11.84	0.13	.90				
Slope					23.05	1.15	20.06	<.001
CMP	−0.816	2.08	−0.39	.70				
Gender	0.25	2.08	0.12	.91				
African American	−3.67	2.08	−1.76	.08				
Hispanic	−1.4	2.43	−0.58	.56				
Caucasian	6.88	2.88	2.39	.02	7.83	2.89	2.71	.01
Asian	13.38	4.23	3.17	<.001	13.18	3.17	4.16	<.001
Concept	1.70	1.52	1.12	.26				
Procedure	0.84	1.69	0.50	.62				

4.3.4. Equation solving tasks

As Table 12 shows, CMP students scored less than non-CMP students in the fall of 2005 ($t = -8.55$, $p < .001$). As for ethnicity, African American students scored less than non-African American students ($t = -5.86$, $p < .001$) whereas Caucasian students score higher than non-Caucasian students ($t = 7.00$, $p < .001$) in the fall of 2005. There were no significant differences between the growth rates of CMP and non-CMP students, between male and female students, between African American and non-African American students, or between Hispanic and non-Hispanic students. Also, the amount of instructional emphasis placed on conceptual understanding did not have a significant impact on the growth rate; nor did the amount of instructional emphasis placed on procedures.

The simple conditional model does suggest that Caucasian students had significantly more improvement than non-Caucasian students ($t = 2.39$, $p < .05$). Asian students also had significantly more growth than non-Asian students ($t = 3.17$, $p < .001$).

The complete conditional model suggests that the estimated mean score for non-CMP, non-African American, and non-Caucasian students was 501.62 and that the estimated mean annual growth rate for non-Caucasian and non-Asian students was 23.05. These estimated mean score and annual growth rate were significantly different from zero ($t = 51.82$, $p < .001$ and $t = 2.71$, $p = .01$, respectively). Caucasian and Asian students had significantly higher growth rates than students outside their ethnic groups. (As Table 12 shows, the t-values and significance levels were $t = 2.71$ ($p < .01$) for Caucasian students, and $t = 4.16$ ($p < .001$) for Asian students.) Specifically, Caucasian and Asian students' annual increases were 23.05 + 7.83 = 30.88 and 23.05 + 13.18 = 36.23 scale points, respectively. The magnitude of effect of this complete model was 16%.

Additional analysis of students of different ethnic groups within the CMP curriculum (not shown in Table 12) revealed that the growth rate of African American students in the CMP program was significantly lower than that of other students ($t = -2.22$, $p < .05$). In addition, Caucasian students in the CMP program benefited more than other students since their growth rate was significantly higher than that of non-Caucasians ($t = 3.63$, $p < .01$). The growth rates of Hispanic and Asian students in the CMP program were not found to differ significantly from students outside their ethnic groups, indicating that the CMP program did not have a biased/different impact on these students' learning rates.

5. Discussion

In the previous section, we presented the results of a longitudinal examination of the CMP and non-CMP students' learning of algebra using four outcome measures (open-ended tasks, translation tasks, computation tasks, and equation solving tasks) and two types of analysis (repeated measures ANOVA and growth curve modeling). In this section, we summarize and discuss these results, focusing, in turn, on our two research questions and on how a procedural or conceptual emphasis in instruction differentially altered the effects on learning of the curricula.

5.1. Curricular effect on basic skills and higher order thinking skills

To reiterate, our first research question was "Do conceptual understanding and higher order thinking skills come at the expense of basic mathematics skills for students who are taught using a Standards-based mathematics curriculum (like CMP)?" Our analyses using repeated measures ANOVA and growth curve modeling showed that on open-ended tasks and translation tasks, CMP students' growth rates were significantly higher than those of non-CMP students, whereas there were similar growth rates for CMP and non-CMP students on computation and equation-solving tasks. Furthermore, the CMP students' growth rates on the open-ended tasks and on the translation tasks were still significantly higher than that of the non-CMP students when students' ethnicity was controlled. These findings suggest that the use of the CMP curriculum is associated with a significantly greater gain in conceptual understanding than is associated with the use of the non-CMP curricula. Furthermore, we found that these relatively greater conceptual gains do not come at the cost of lower basic skills, as evidenced by the comparable gains attained by CMP and non-CMP students on the computation and equation solving tasks.

So, why do CMP students show significantly greater growth on the conceptually oriented measures (open-ended and translation tasks) than non-CMP students? One interpretation of the data is that the CMP students' significantly greater gains on both open-ended tasks and translation tasks are related to the nature of the curriculum. Even a cursory comparison of the CMP curriculum with the non-CMP curricula used in our study reveals major differences between them. Nonetheless, we did an in-depth comparison of the approaches to algebra taken by the two types of curricula. We found major differences in the development of fundamental algebraic ideas, and these differences appear to be related to the differences we observed (Cai et al., 2010; Nie et al., 2009). For instance, the CMP curriculum uses a functional approach to introduce variables and equations, whereas the non-CMP curriculum uses a structural approach to introduce variables and equations. By way of example, in the 7th grade, when the CMP curriculum introduces a formal symbolic procedure for solving equations, the focus is on making sense of the procedure by providing a contextual meaning for each step of the equation solving process (see Table 1). The non-CMP curricula take a much different approach, typically introducing the formal symbolic procedure for solving equations by illustrating and employing the additive property of equality on "naked" equations.

Based on differences like this, which we found in our analysis of the CMP and non-CMP curricula, it is likely that features of the CMP curriculum contributed to the CMP students' significantly greater gains on both the open-ended and translation tasks. It might be expected that a similar advantage would accrue in the equation solving of students who used the non-CMP curricula. However, even though the non-CMP curricula include many more de-contextualized equation-solving exercises than the CMP curriculum, our data did not show greater achievement gains for the non-CMP students on the equation-solving tasks, which were also de-contextualized.

5.2. Curricular effect on students of color

Our second research question is "*How does the use of a Standards-based curriculum impact the learning of students of color as compared to Caucasian students*?" the analyses using growth curve model showed that African Americans who use the CMP curriculum had a smaller growth rate than the other ethnic groups on the computation and equation-solving tasks. It also showed that, on the open-ended and translation tasks, the African American and Hispanic students had at least as large a growth rate as the Caucasians.

In summary, the answer to our second research question is "no" for the conceptual-based measures we employed and "yes" for the procedural-based measures. Analyses using both repeated measures ANOVA and growth curve modeling showed that the use of a *Standards*-based curriculum like CMP improves the mathematics achievement for all students. Moreover the CMP program contributes to significantly higher growth than the non-CMP programs for all ethnic groups on both open-ended tasks and translation tasks (especially on the open-ended tasks). However, the findings do suggest that the use of the CMP program has a negative impact on African American students' growth on both the computation and equation-solving tasks when compared to other ethnic groups using CMP and when compared to African Americans using a non-CMP curriculum. On the other hand, the use of the CMP program has a positive impact on Hispanic students' growth on both computation and equation-solving tasks. These findings address the call for more research on the equity issues of curriculum for students of color (Lubienski & Gutiérrez, 2008; Schoenfeld, 2002). Further analysis is needed to understand why the impact of curriculum on African Americans' achievement gains was different than the impact of curriculum on Hispanic Americans' achievement gains over the three middle school years.

5.3. Conceptual and procedural emphases in the classroom

One of the major contributions of this study is its longitudinal investigation of how classroom instruction that emphasizes procedures or concepts influences the effect of the curriculum on student learning. However, our growth curve modeling analysis did not detect a statistically significant impact of either a conceptual or a procedural emphasis on the growth rate of any of our outcome measures except for translation component. Significantly, we found that the relatively greater growth rates attributable to the use of the CMP curriculum disappeared when the conceptual emphasis in the classroom was controlled. That is, when the emphasis on conceptual understanding was the same in both CMP and non-CMP classrooms, then the rates of growth on the translation tasks were also the same. Related to this outcome, we also found that the level of conceptual emphasis in the classroom could significantly predict the growth rate of the students on the translation component of the multiple-choice tasks.

One possible reason that the analyses using growth curve modeling was unable to use the level of conceptual or procedural emphasis to predict achievement gains on all of the outcome measures may be related to the small variances in the levels of conceptual and procedural emphases in both CMP and non-CMP classrooms. For example, the variance in conceptual emphasis within CMP classrooms was 0.94 and that within non-CMP classrooms was 0.64.

While CMP teachers emphasize significantly more conceptual understanding than non-CMP teachers, the overall average level of conceptual emphasis in CMP classrooms was rated only slightly above a score of 3 on a five-point scale. Similarly, while non-CMP teachers emphasize significantly more procedural understanding than CMP teachers, the overall average level of procedural emphasis in non-CMP classrooms is only slightly above a score of 3 on a five-point scale. Even though the differences are statistically significant between CMP and non-CMP instruction, the small absolute differences in both conceptual and procedural emphases may not be sensitive enough to accurately predict achievement gains for both CMP and non-CMP students. Nevertheless, the measures of conceptual and procedural emphases used in the LieCal Project provide a new way to characterize classroom instruction.

5.4. Future analyses

In this article, we analyzed data collected using a longitudinal design across the three middle school years to examine the differential impact of a reform or a traditional curricula on algebra-related learning of students with different ethnic backgrounds. We did not, however, examine the growth of students' learning from one grade level to another or at each grade level. A future direction of the analysis will be to determine the impact of curriculum type on pairs of successive grade levels and within each grade level.

Another direction of study in the future will be to analyze how the type of curriculum used in middle school impacts students' transition to the 9th grade algebra classes in high school. In fact, we have already collected the data needed to perform that analysis for the students who participated in the grades 6–8 portion of the LieCal project. The analysis will be interesting because all the former middle school CMP and non-CMP students used the same mathematics curriculum in the 9th grade. Further, this type of analysis is significant because it examines something that has not been studied by other researchers, namely the impact of curriculum on learning across grade bands.

Acknowledgements

This article is based on a paper presented at the Annual Meeting of the American Educational Research Association. Denver, CO, April 29–May 3, 2010.

The research reported here is supported by grants from the National Science Foundation (ESI-0454739 and DRL-1008536). Any opinions expressed herein are those of the authors and do not necessarily represent the views of the National Science Foundation.

Appendix A

Sample tasks on the LieCal assessment

Sample multiple-choice tasks

Translation

Which number sentence is correct?

One pound of shrimp costs \$3.50 more than one pound of fish.

a. shrimp's cost per pound = fish's cost per pound + \$3.50
b. shrimp's cost per pound + \$3.50 = fish's cost per pound
c. shrimp's cost per pound + fish's cost per pound = \$3.50
d. shrimp's cost per pound = fish's cost per pound – \$3.50

Integration

Which numbers are needed to solve this problem?

Bill has \$2000 in his savings account. He paid a \$195 registration fee, a \$100 student health fee, and a \$50 activity fee. How much money did he spend?

a. 2000, 195, 100, 50
b. 195, 100, 50
c. 100, 50
d. 50

Planning

Which operations could you carry out to solve this problem?

Fifteen pencils come in each box at the store. You buy 3 boxes on Monday, 2 boxes on Wednesday, and 1 box on Friday. How many pencils did you buy during this period?

a. add, then multiply
b. add, then divide
c. subtract only
d. divide only

Computation
$14.04 \div 13 =$
a. 1.08
b. 1.8
c. 10.8
d. 18

Solving equations
Find the value of x so that $x - 5 = 5$
a. 0
b. 1
c. 10
d. 25

A sample open-ended task

Sally is having a party
The first time the doorbell rings, 1 guest enters
The second time the doorbell rings, 3 guests enter
The third time the doorbell rings, 5 guests enter
The fourth time the doorbell rings, 7 guests enter
Keep going in the same way. On the next ring, a group enters that has 2 more persons than the group that entered on the previous ring
A. How many guests will enter on the 10th ring? Explain or show how you found your answer
B. How many guests will enter on the 100th ring? Explain or show how you found your answer
C. 299 guests entered on one of the rings. What ring was it? Explain or show how you found your answer
D. Write a rule or describe in words how to find the number of guests that entered on each ring

References

Bailey, R., Day, R., Frey, P., Howard, A. C., Hutchens, D. T., McClain, K., et al. (2006a). *Mathematics: Applications and concepts (teacher wraparound edition), course 1*. Columbus, OH: The McGraw-Hill Company.
Bailey, R., Day, R., Frey, P., Howard, A. C., Hutchens, D. T., McClain, K., et al. (2006b). *Mathematics: applications and concepts (teacher wraparound edition), course 2*. Columbus, OH: The McGraw-Hill Company.
Bailey, R., Day, R., Frey, P., Howard, A. C., Hutchens, D. T., McClain, K., et al. (2006c). *Mathematics: applications and concepts (teacher wraparound edition), course 3*. Columbus, OH: The McGraw-Hill Company.
Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher 25(9) 6-8, 14.*
Bednarz, N., Kieran, C., & Lee, L. (Eds.). (1996). *Approaches to algebra: perspectives for research and teaching*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
Cai, J. (2003). What research tells us about teaching mathematics through problem solving. In F. K. Lester, Jr., (Ed.). *Teaching mathematics through problem solving: Prekindergarten-grade 6* (pp. 241–254). Reston, VA: National Council of Teachers of Mathematics.
Cai, J. (2000). Mathematical thinking involved in U.S. and Chinese students' solving process-constrained and process-open problems. *Mathematical Thinking and Learning: An International Journal, 2*, 309–340.
Cai, J., & Knuth, E. (2011). *Early algebraization: A global dialogue from multiple perspectives*. New York, NY: Springer.
Cai, J., Nie, B., & Moyer, J. (2010). The teaching of equation solving: Approaches in *standards*-based and traditional curricula in the United States. *Pedagogies: An International Journal, 5*(3), 170–186.
Cai, J., Wang, N., Moyer, J.C., Nie, B., & Wang, C. (submitted). Learning mathematics from classroom instruction using Standards-based and traditional curricula: A longitudinal study, Submitted for publication.
Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Arithmetic and algebra in elementary school*. Portsmouth, NH: Heinemann.
Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443–494.
College Board. (2000). *EQUITY 2000: A systemic education reform model*. Washington DC: Author.
Compton, D. L. (2000). Modeling the growth of decoding skills in first grade children. *Scientific Studies in Reading, 4*, 219–259.

Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J., & Jones, D. (2009). Fostering high-quality teaching with an enriched curriculum and professional development support: The head start REDI program. *American Educational Research Journal, 46*, 567–597.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Press.

Harwell, M. R., Post, T. R., Maeda, Y., Davis, J. D., Cutler, A. L., Anderson, E., et al. (2007). Standards-based mathematics curricula and secondary students' performance on standardized achievement tests. *Journal for Research in Mathematics Education, 38*, 71–101.

Herman, R., Boruch, R., Powell, R., Fleischman, S., & Maynard, R. (2006). Overcoming the challenges: A response to Alan H. Schoenfeld's "What doesn't work". *Educational Researcher, 35*(2), 22–23.

Howson, G., Keitel, C., & Kilpatrick, J. (1981). *Curriculum development in mathematics*. Cambridge; New York: Cambridge University Press.

Kaput, J. (1999). Teaching and learning a new algebra. In E. Fennema & T. A. Romberg (Eds.), *Mathematical classrooms that promote understanding* (pp. 133–155). Mahwah, NJ: Lawrence Erlbaum Associates.

Katz, V.J. (2007). *Algebra: Gateway to a technological future*. Retrieved August 17, 2010 from http://www.maa.org/algebra-report/Algebra-Gateway-Tech-Future.pdf.

Kieran, C. (1997). Mathematical concepts at the secondary school level: the learning of algebra and functions. In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 133–158). East Sussex: Psychology Press.

Kilpatrick, J. (2003). What works? In S. L. Senk & D. R. Thompson (Eds.), *Standards-based school mathematics curricula: What are they? What do students learn?* (pp. 471–488). Mahwah, N.J Lawrence Erlbaum Associates.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Newbury Park, CA: Sage.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (2002a). *Connected mathematics*. Upper Saddle River, NJ: Prentice Hall.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillips, E. D. (2002b). *Moving straight ahead*. Upper Saddle River, NJ: Prentice Hall.

Lane, S., Silver, E. A., Ankenmann, R. D., Cai, J., Finseth, C., Liu, M., et al. (1995). *QUASAR cognitive assessment instrument (QCAI)*. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates Inc.

Loveless, T. (2008). *The misplaced math student. The 2008 Brown Center Report on American Education: How well are American students learning?* Washington: Brookings.

Lubienski, S. T., & Gutiérrez, R. (2008). Bridging the gaps in perspectives on equity in mathematics education. *Journal for Research in Mathematics Education, 39*, 365–371.

Mathematical Sciences Education Board. (1998). *The nature and role of algebra in the K-14 curriculum: proceedings of a national symposium*. Washington D.C. National Research Council.

Mayer, R. E. (1987). *Educational psychology: A cognitive approach*. Boston: Little & Brown.

Mislevy, R. J. (1995). What can we learn from international assessments? *Educational Evaluation and Policy Analysis, 17*(4), 419–437.

National Academy of Education. (1999, March). Recommendations regarding research priorities: An advisory report to the National Educational Research Policy and Priorities Board. Retrieved August 17, 2010 from http://www.naeducation.org/Research_Priorities_Publication.pdf.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, D.C. U.S. Department of Education.

National Research Council. (2001). In J. Pelligrino, N. Chudowsky, & R. Glaser (Eds.), *Knowing what students know: The science and design of educational assessment. Committee on the Foundations of Assessment*. Washington, DC: National Academy Press.

National Research Council. (2004). In J. Confrey & V. Stohl (Eds.), *On evaluating curricular effectiveness: Judging the quality of k-12 mathematics evaluations*. Washington, DC: National Academy Press.

Nie, B., Cai, J., & Moyer, J. C. (2009). How a standards-based mathematics curriculum differs from a traditional curriculum: With a focus on intended treatments of the ideas of variable. *ZDM (International Journal on Mathematics Education), 41*(6), 777–792.

Pelavin, S. H., & Kane, M. (1990). *Changing the odds: Factors increasing access to college*. New York: College Board.

Post, T. R., Harwell, M. R., Davis, J. D., Maeda, Y., Cutler, A., Andersen, E., et al. (2008). Standards-based mathematics curricula and middle-grades students' performance on standardized achievement tests. *Journal for Research in Mathematics Education, 39*, 184–212.

RAND Mathematics Study Panel (2003). Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education. MR-1643-OERI.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S., Bryk, A., & Congdon, R. (2005). *HLM 6.03 for Windows*. Multivariate Software.

Reardon, S. F., & Galindo, C. (2009). The hispanic-white achievement gap in math and reading in the elementary grades. *American Research Journal, 46*, 853–891.

Reys, R., Reys, B., Lapan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standards-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education, 34*, 74–95.

Riordan, J., & Noyce, P. (2001). The impact of two standards-based mathematics curriculum in student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32*, 368–398.

Rivette, K., Grant, Y., Ludema, H., & Rickard, A. (2003). *Connected mathematics project: Research and evaluation summary 2003 edition*. Upper Saddle River, NJ: Pearson Prentice Hall.

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education, 34*, 228–259.

Schoenfeld, A. H. (2006). What doesn't work: The challenge and failure of the works clearinghouse to conduct meaningful reviews of studies of mathematics curricula. *Educational Researcher, 35*(2), 13–21.

Show-Me Center (2002). *CMP implementation map.* http://www.math.msu.edu/cmp/Overview/ImplementationMap.htm.

Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher, 31*, 13–25.

Senk, S. L., & Thompson, D. R. (Eds.). (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Mahwah, N.J Lawrence Erlbaum Associates.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50–80.

Tarr, J. E., Reys, R. E., Reys, B. J., Chávez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39*, 247–280.

Usiskin, Z. (1999). Which curriculum is best? *UCSMP Newsletter, 24*(Winter), 3–10.

Wilson, S. M., & Floden, R. E. (2001). Hedging bets: standards-based reform in classrooms. In S. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states (One hundredth yearbook of the National Society for the Study of Education, Part 2, pp. 193–216*. Chicago: University of Chicago Press.

Wu, H. (1997). The mathematics education reform: Why you should be concerned and what you can do. *American Mathematical Monthly, 104*, 946–954.