

ADDRESSING MATHEMATICAL MISCONCEPTIONS:
IS PROBABILITY THE INDEPENDENT VARIABLE?

Christopher Rakes
University of Louisville

Author Note: This paper was part of a preliminary dissertation draft. Please refer to Rakes (2010)
for the final dissertation draft (UMI No. 3415205)

A paper presented at the annual meeting of the American Educational Research Association,
Research in Mathematics Education Special Interest Group, Denver, CO

April, 2010

ABSTRACT

This study examined the impact of probability instruction on misconceptions that are common to rational numbers, algebra, geometry, and probability and the impact of orientation, discernment, and context on the development of mathematical misconceptions using a mixed methods research design. Five foundational concepts were identified through a review of literature that, if not addressed, create persistent difficulties for student mathematics learning: rational number meaning, additive/multiplicative structures, absolute/relative comparison, variable meaning, and spatial reasoning misconceptions. The intervention probability unit lasted approximately two weeks.

This study used mixed methodology to analyze data from a randomly assigned sample of students from an untreated control group design with a switching replication. Document analysis was used to examine patterns in student responses to items on the mathematics knowledge test. Multiple imputation was used to reduce selection bias from missing data. Item response theory was used to compute item difficulty, item discrimination, and item guessing coefficients. Generalized hierarchical linear modeling was used to explore the impact of item, student, and classroom characteristics on errors due to misconceptions.

These analyses resulted in 4 key findings. (1) Mathematics misconception errors often appear as procedural errors. (2) A classroom environment that fosters enjoyment of mathematics and value of mathematics are associated with reduced misconception errors. (3) Higher mathematics self confidence and motivation to learn mathematics is associated with reduced misconception errors. (4) Probability instruction may not affect misconceptions directly, but it

may help students develop skills needed to bypass misconceptions when solving difficult problems.

DRAFT

ADDRESSING MATHEMATICAL MISCONCEPTIONS: IS PROBABILITY THE INDEPENDENT VARIABLE?

Transitioning from whole numbers to rational numbers is often one of the key stumbling blocks preventing students from learning mathematics deeply in elementary and middle school (Moss, 2005). These problems persist throughout high school, and add to the difficulties of transitioning from arithmetic to algebra (Kilpatrick, Swafford, & Findell, 2001). Throughout these transition periods, students may attempt to incorporate new information into their current knowledge base without sufficient understanding to successfully bridge the ideas (MacGregor and Stacey, 1997). The conceptual errors in thinking that can result from such unsuccessful attempts lead to a set of misconceptions that inhibit learning in every strand of mathematics, especially rational numbers, algebra, geometry, and probability. These underlying misconceptions may be difficult to correct because they are, in turn, merely symptoms of deeper root causes.

Several researchers have proposed that probability instruction may hold the key to reducing these common misconceptions because of its characteristic abundance of concrete simulations and experiments (e.g., Agnoli, 1987; Agnoli & Krantz, 1989; Bar-Hillel & Falk, 1982; Falk, 1992; Falk & Lann, 2008; Freudenthal, 1970, 1973, 1983; Shaughnessy, 1992; Shaughnessy & Bergman, 1993; Watson & Shaughnessy, 2004). These concrete applications may help students successfully develop the relational understanding that is necessary for deep mathematical learning. The purpose of this study is to investigate misconceptions common to rational numbers, probability, algebra, and geometry by examining two research questions:

- 1) Does probability instruction reduce critical misconceptions in probability, rational numbers, algebra, or geometry?

- 2) Do student attitudes toward mathematics influence the emergence of errors due to misconceptions on mathematical tasks?
- 3) Does student metacognition influence the emergence of errors due to misconceptions on mathematical tasks?

Rational Number Misconceptions

Rational numbers confound student conceptual understanding in mathematics more than whole numbers, in part because of the multiple uses of rational numbers and ways to represent them (Fuson, Kalchman, & Bransford, 2005; Kilpatrick et al., 2001; Lamon, 2007; Moss, 2005). Two of the most common rational number relationships have been described as “part/part” and “part/whole” (e.g., Baturu, 1994; Behr, Harel, Post, & Lesh, 1992). The part/whole relationship may be the more important of the two relationships for understanding fractions (Behr et al., 1992; Fosnot & Dolk, 2002).

Misconceptions in Algebra and Geometry

Students beginning the study of algebra face a transitional barrier resulting from the structural and abstract characteristics of algebraic concepts (Carragher and Schliemann, 2007; Howe, 2005; Kieran, 1989, 1992; Vogel, 2008). Second, the learning of algebra requires students to learn a language of mathematical symbols that is completely foreign to their previous experiences (Blanco & Garrote, 2007; Kilpatrick et al., 2001; Socas Robayna, 1997). In some cases, students are completely unaware that any meaning was intended for the symbols (Küchemann, 1978). In other cases, they may know that meaning exists, but limited understanding prevents them from ascribing meaning to the symbols, or they may assign erroneous meaning to the symbols (Küchemann, 1978).

Student orientation toward geometry is quite different from that of algebra: Students are

often intrinsically motivated to study the properties that govern the shapes encountered in daily life (Engel, 1970; Freudenthal, 1973). In spite of this motivational factor, students still struggle with errors and misunderstandings in geometry due to limited spatial reasoning (Clements & Battista, 1992). The van Hiele (1959/1984a) five-level framework is especially helpful for illuminating student spatial reasoning processes and how they distinguish between objects and representations.

The Advantage of Probability

Stone, Alfeld, and Pearson (2008) echoed the sentiments of Freudenthal (1970): In order to guide students to deep mathematical learning, mathematical content must be tied to authentic experiences to which students can relate. Probability offers such a connection between mathematics and the real world naturally (Liu & Thompson, 2007), and its de-emphasis in U.S. high school mathematics curricula may account for many of the difficulties students have connecting abstract mathematical ideas to concrete examples (Davis, 1992). In spite of the ability of probability to bridge the gulf between the abstract and concrete, several reasons explain its exclusion from mathematics curricula. First, teachers are typically less familiar with probability content than other areas of mathematics (Jendraszek, 2008; Swenson, 1998). Compounding this problem is the fact that probability is often viewed as a second-rate topic (Mitchell, 1990; Shaughnessy, 2006).

A Conceptual Framework to Model Mathematics Learning

The similarity between reasoning across mathematics content areas suggests a pathway of learning that either results in understanding or misconceptions and errors. To develop a model that traces these pathways, several factors must be considered. First, the introduction of new concepts is typically accompanied with tasks or problems for the students to complete. The

characteristics of these tasks (e.g., task complexity, difficulty, discrimination between ability levels) may influence how students interpret the new material. Second, students must filter tasks through their own knowledge framework. Third, the pedagogical emphasis on either concepts or procedures direct students to develop either relational or instrumental understanding (Skemp, 1976/2006). If students learn relationally, then the conceptual understanding they develop may produce stronger, more consistent procedural skills, which in turn may reinforce deeper more robust conceptual understanding. This understanding may then be integrated into a student's knowledge framework for use with future tasks.

Alternatively, the development of instrumental understanding leads to the development of procedures without meaning, with incomplete or erroneous meaning, or even the lack of awareness of meaning (Skemp 1976/2006). Misunderstanding the meaning of mathematical objects in some way is the very essence of misconceptions. Mathematical misconceptions result in errors that are often difficult for teachers to prevent or obstruct. Researchers have repeatedly found that systematic errors due to misconceptions rather than faulty reasoning adhere to patterns of over- or under-generalization of properties or concepts for a particular task (e.g., Chang, 2002; Falk, 1992; Fuys & Liebov, 1997; Kalchman & Koedinger, 2005; Van Dooren, De Bock, Depaepe, Janssens, and Verschaffel, 2003). For example, in geometry, Fuys and Liebov (1997) suggested that students struggling to move from a visualization level of spatial reasoning to an analysis level may under-generalize geometric properties by including irrelevant characteristics of a shape in their mental framework; or conversely, they may over-generalize relationships between figures by discarding any number of a shape's unique properties. If unchecked, these misconceptions may be integrated into students' mathematical understanding, thereby influencing future learning.

Difficulties Inherent to Addressing Mathematical Misconceptions Directly

Multiple attempts to develop interventions for reducing misconceptions have met with limited success. Some of these efforts have focused on addressing task-specific errors (e.g., Rosnick & Clement, 1979). One difficulty with such a strategy is that if the error was due to a misconception, the underlying misconception will remain in the student's knowledge framework to adapt and reappear in the same or other task. Other endeavors have attempted to address the reasoning that leads to an error using a variety of strategies such as worked examples (e.g., Fisher, 1988; Phillippe, 1992; Rosnick & Clement, 1980). Directly addressing erroneous reasoning appeared to make no significant improvement in student learning (Weinberg, 2007).

Weinberg (2007) suggested that another reason student errors can be so insidious is that students attempt to adapt their knowledge base to the problem scenario, sometimes accurately and sometimes not. The adaptive nature of these errors suggests that the reasoning processes are built on a deeper foundation of understanding relating to the structure and meaning of mathematical ideas (Kieran, 2007, 2008, 2009).

An even more robust intervention design may be needed to alter students' mathematical thinking and reasoning. Rather than targeting reasoning processes directly, such an intervention might focus instead on transforming the instrumental understanding responsible for difficulties in meaning that can lead to misconceptions into relational understanding. If a teaching intervention targets the development of meaning and connections, then misconceptions that develop may be only a normal, temporary part of the learning process (Resnick, 1983).

Mathematical Task Characteristics

Teachers typically introduce new concepts by presenting a task or problem as a motivation for learning the mathematical concept. Rousseau (1976) identified eight task

characteristics that potentially influence how students internalize the meaning of the task and its connection to the underlying concept: task identity, task autonomy, skill variety, task variety, task feedback, task learning, dealing with others, and task significance. Task identity refers to the ownership a student assumes for an activity. Task autonomy, closely aligned with identity, focuses on the degree of independence students have in decision making throughout a task. Skill variety emphasizes the breadth and depth of skills required to complete a particular task. Task variety, on the other hand, refers to the breadth of subjects and courses provided by a school. Task learning represents the breadth and scope of opportunities for obtaining new skills, what Hiebert and Grouws (2007) referred to as opportunity to learn. Task feedback speaks to the amount of feedback students receive from a task versus the feedback from teachers. Catanzaro (1997) maintained that task feedback creates a more stimulating, positive learning environment over instructor feedback. Rousseau (1976) defined dealing with others as “the opportunity to interact with teachers, teaching assistants and other faculty” (p. 3). Many researchers would also emphasize the importance of interactions with other students (e.g., Berg, 1993; Freeman, 1997; Henderson & Landesman, 1995; Nichols & Miller, 1994; Parham, 1993; Slavin & Karweit, 1982; Slavin & Lake, 2008; Slavin, Lake, & Groff, 2009; Whicker, Bol, & Nunnery, 1997). Task significance represents student perceptions of a particular task’s relevance to life beyond academic concerns. Rousseau (1976) found that task significance may have the strongest impact of her eight task characteristics.

Student Thought Processes Influencing Mathematical Misconceptions

Erroneous thinking resulting from misconceptions is often stable and robust, interfering with a student’s ability to learn mathematics (Moschkovich, 1998). Researchers tend to agree that a possible key to addressing these issues may lie in the alignment of student thought

processes with mathematical logic (e.g., Behr, 1980; Blanco & Garrote, 2007; Collis, 1975; Enfedaque, 1990; Kieran, 1980; Palarea Medina, 1999; Socas Robayna, 1997) and the connection of specific misconceptions to the student's larger knowledge framework (e.g., Moschkovich, 1998; Smith, diSessa, & Roschelle, 1993). This knowledge framework includes (at least) four components that can influence whether a student develops relational or instrumental understanding: (1) Discernment; (2) Orientation toward mathematics; (3) Individual context; and (4) Environmental Context.

Discernment. Discernment has been defined as an aspect of knowledge that encompasses the active, cognitive components of learning (Ronau & Rakes, 2010; Ronau, Rakes, Wagener, & Dougherty, 2009; Ronau, Wagener, & Rakes, 2009). Kant (1786/1901) proposed that cognition is engaged through the process of perceptions leading to conceptions, which in turn lead to ideas. Davis (1992), comparing Japanese to American tests, considered the influence of such perceptions to be paramount to deep mathematical learning:

Perhaps 75 one-step problems on a test will produce about the same ranking of students as will 6 multistep problems that require serious thought (and perhaps some originality). *But the message that they send to students is entirely different.* The one-step problems say to students, "You do not have to do much hard thinking in mathematics, nor must you be very creative; all you have to do is pay attention in class, memorize dutifully, practice diligently, and you will get no surprises on the tests." The Japanese tests send a different message — rather more in the spirit of the contest problems that a very few U.S. students encounter — where it is more clear from the outset that, if you have developed nothing more than routine skills, you will be hopelessly ineffective. You *must* strive for

ingenuity and originality (p. 725).

Davis (1992) went on to consider the meaning of mathematics from a cognitive perspective. He gave three examples of problems whose solution required the addition of whole numbers. These problems differed in the degree of decision making required about each contextual situation prior to concluding that addition is needed for each.

Now, here is the main point behind these three examples: Most people who have not had an opportunity to think seriously about such matters would claim that the mathematics is that part of the problem that the calculator did. They might find the decisions...or the choice of arithmetical operations...to be thought provoking, but they would probably not consider them an essential part of the mathematics...they might not even notice that there was any thinking involved other than the computation that the calculator carried out. I would argue that such observers are precisely wrong. There is very little mathematics in the actual carrying out of the computations...The mathematics lies mainly in analyzing the real situation and deciding how to represent it in an appropriate abstract symbolic form (Davis, 1992, p. 727).

Schoenfeld (1992) agreed with Davis' conceptualization of the nature of mathematical learning. He added that mathematical problem solving requires a great deal of metacognitive regulation and that such behavior is learned best through "domain-specific instruction" (p. 357). In an earlier work (Schoenfeld, 1982), he considered three types of analysis to be important to mathematical problem solving: analysis of tactical knowledge (i.e., domain-specific facts and procedures), analysis of control knowledge (i.e., strategic/executive behavior), and analysis of belief systems. The analysis of control knowledge speaks directly to metacognition, the

regulation of cognitive processes. Several other researchers have suggested that cognitive and meta-cognitive skills filter student ability to understand mathematical concepts (Andrade & Valtcheva, 2009; Dermitzaki, Leondari, & Goudas, 2009; Fuson et al., 2005; Lin, Schwartz, & Hatano, 2005; Nemirovsky & Ferrara, 2009; Usher, 2009). Swanson (1990) found that the development of metacognition may operate independently of aptitude and may impact learning more:

On the surface, it appears that high metacognitive skills can compensate for overall ability by providing a certain knowledge about cognition. This knowledge allows low-aptitude/high-metacognitive children to perform in ways similar to those of children with high aptitude. Thus, one may argue that measures of metacognition and general aptitude in the present study are tapping different forms of knowledge, and that high performance on the problem-solving tasks is more closely related to higher performance on the metacognitive measures than on the aptitude measures (Swanson, 1990, p. 312).

Schraw and Dennison (1994) identified two constructs that measure metacognition: knowledge of cognition and regulation of cognition. They found that, although the two constructs are correlated, each may affect cognitive performance in a unique way. Other studies have shown that students use these cognitively-based discernment faculties to connect abstract concepts to concrete representations (e.g., Secada, 1992; Spillane, 2000; Von Minden, Walls, & Nardi, 1998).

Orientation toward mathematics. Schoenfeld's (1982) third type of analysis focused on student beliefs. He posited that student beliefs about the nature of a mathematical task can greatly influence the degree of cognitive effort expended for the task. Schoenfeld (1985)

conducted a survey of 230 students in three high schools. He found three aspects to student beliefs about mathematics. (1) Students in his sample attributed success in mathematics to work rather than luck. (2) Students in his sample disagreed that mathematics solutions were either “right” or “wrong.” They also declared the importance of teaching multiple ways to solve mathematics problems. This response surprised Schoenfeld because “very little of such teacher behavior was observed in the classroom studies...their response suggests either a strong acceptance of the mythology about teaching, or some strong degree of wishful thinking” (p. 14). (3) Students view mathematics learning as largely dependent on memorization while simultaneously viewing it as a means to develop logical thinking.

McLeod (1992) agreed with Schoenfeld’s description of beliefs and attitudes as components of affect; however, he added a third, distinct component category: emotions. Emotional reactions to mathematics learning occur when students experience obstacles to solutions. Such obstacles elicit negative feelings such as tension, frustration, fear, anxiety, embarrassment, and panic. Once obstacles are overcome, positive emotions return. He maintained that one goal of mathematics pedagogy should be to reduce the occurrence of these negative emotions. From attitudes, beliefs, and attitudes, seven subconstructs of affect emerge: confidence, self concept, self efficacy, anxiety, effort and ability attributions, learned helplessness, and motivation.

Schoenfeld and McLeod agreed that affect and cognition are linked (Schoenfeld, 1989; McLeod, 1992). Schoenfeld (1989) found that beliefs and attitudes influence the way people develop conceptions about mathematics, directly and indirectly impacting their mathematical ability. Barkatsas, Kasimatis, and Gialamas (2009) found that high levels of mathematics achievement are associated with positive attitudes toward learning mathematics; positive

attitudes, in turn, are associated with mathematics confidence and affective engagement. Ismail (2009) found that self-confidence appeared to supersede the impact of socio-economic disadvantage on student achievement.

In summary, components of affect such as beliefs, attitudes, and emotions mold student orientations toward mathematics. Pedagogical strategies within mathematics influence the development of conceptions or misconceptions as a result of their attention to orientation.

Individual context. The contextual factors that students bring to a mathematical learning situation interact in multiple ways to influence how students interpret mathematical concepts. These individual context factors refer to characteristics such as gender, race, culture, socio-economic status, parent education levels, background experiences, and learning styles (Ronau et al., 2009; Ronau & Rakes, 2010; Ronau, Wagener, & Rakes, 2009).

Evidence has suggested that boys and girls construct their understanding of mathematics differently (Fennema & Sherman, 1977) and hold different attitudes toward mathematics (Sherman & Fennema, 1978). Although moderate changes have occurred over time, inequity between genders still exists (Carrell, Page, & West, 2009; Fennema, 2000; Mendick, 2008; Van Langen, Rekers-Mombarg, & Dekkers, 2008; Wei & Hendrix, 2009; Zohar & Gershikov, 2008).

Kozol (1992, 2005) examined educational practices across the country and asserted that inequalities also continue to exist across racial lines. Snipes and Waters (2005) agreed with Kozol's assessment, conducting a case study in a single state. Lubienski (2001) and Lim (2008) found that race and class interact to produce an effect on mathematics achievement. Class measures include factors such as parent education levels and socio-economic status (SES). Parent education levels, one measure of SES, significantly predicted above average achievement during the Third International Math and Science Study (TIMSS; Schreiber, 2000). Lehrer, Strom,

and Confrey (2002) found that prior mathematical experiences influence student orientation toward mathematics. Anderson (1990) asserted that cultural influences overshadow gender and racial effects on equity in student achievement. Nelson, Joseph, & Williams (1993) agreed with Anderson, claiming that culture also has a direct bearing on affect. Strutchens (1995) proposed the use of a five-dimensional framework for increasing equity in mathematics education: content integration, knowledge construction, prejudice reduction, equitable pedagogy, and empowering school and social culture.

Alomar (2007) and Esposito Lamy (2003) linked gender, race, culture, and affect with family variables such as parenting style and poverty. Lopez, Gallimore, Garnier, and Reese (2007) found that for immigrant populations, family factors influence English language literacy, which in turn affects student mathematics achievement.

Personal characteristics such as learning styles, personality, and temperament also influence how students learn mathematics. The Silver-Strong studies (Silver, Brunsting, & Walsh, 2008; Silver, Strong, & Perini, 1997; Strong, Perini, Silver, & Thomas, 2004; Strong, Silver, & Perini, 2001) together with the work of Keirsey (1998) suggest a link between learning styles and personality. Keirsey (1998) described personality in terms of the Myers-Briggs notation. In this framework, a person may be Introverted (I) or extraverted (E); rely more on intuition (N) or the senses (S) to interpret a situation; rely more on feelings (F) or thinking (T) to make decisions; and, prefer routine (J for judgment) or sponteneity (P for perceiving), resulting in 16 different personality styles that he grouped into four categories with internal reliability ratings between 0.82 and 0.83 (Alpine Media Corporation, 2003). Silver et al. (1997) used the same constructs to determine four categories of learning styles: Mastery, Understanding, Interpersonal, and Self-Expressive Learners. The dependence on these two frameworks on the

Myers-Briggs constructs (Myers, 1962) suggests a possible link between learning styles and personality. These frameworks directly map onto one another (Table 1).

Table 1

Alignment of Keirsey Personality Framework with Silver-Strong Learning Styles

Silver-Strong Learning Style	Values and Educational Preferences	Associated Keirsey Personality Types
Mastery	Value: Clarity and Practicality Prefer: procedure, drill and practice, concrete, closed questioning.	Guardian Administrators: ISTJ; ESTJ Artisan Operators: ISTP; ESTP
Understanding	Value: Logic and Evidence Prefer: logic, debate, inquiry, independent study, argumentation, and why questions.	All Rational Subgroups: INTJ; INTP ENTJ; ENTP
Interpersonal	Value: The ability to help others Prefer: topics that affect lives, cooperative/collaborative learning, and teacher attention to successes and struggles.	Guardian Conservators: ISFJ; ESFJ Artisan Entertainers: ISFP; ESFP
Self Expressive	Value: Originality and aesthetics Prefer: use of imagination to explore ideas, creative artistic activity, open-ended questions, and generating possibilities and alternatives.	All Idealist Subgroups: INFJ; INFP ENFJ; ENFP

Understanding the role of values and preferences of the various types of learners directly impacts the equitable teaching of mathematics (Gardner & Hatch, 1989). Second, the traditional mathematics education described by Fey (1979) that still continues today (Hiebert, 2003; Hiebert & Grouws, 2007; Stigler & Hiebert, 1997) targets mastery learners almost exclusively, while they account for only about 35% of the population (Silver et al., 1997). In smaller samples, such as a single high school, the mastery learners have been found to account for far lower percentages (24% in Tungate, 2008). That mathematics teachers tend to be mastery learners themselves seems likely and would account for the disproportionate bent toward traditional practices.

Personality has been framed most prominently as five major constructs known as “The Big Five:” Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Ahadi & Rothbart, 1994, p. 189). Personality emerges from temperament, but assessment of adult personality may not map directly from temperament (Rothbart, Ahadi, &

Evans, 2000). For example, cognitive self-concept may supersede temperamental tendencies (i.e., beliefs about how a person would like to be, should be, and is in reality are difficult to separate).

Posner and Rothbart (2007), Rothbart and Jones (1998), Rueda, Rothbart, Saccomanno, and Posner (2007) and Rudasill (2009) asserted that Attention, one temperament factor, may influence the learning of mathematics both directly and indirectly. “Everywhere in cognitive neuroscience, specific brain networks seem to underlie performance. However, some of those networks have the important property of being able to modify the activity in other networks” (Posner & Rothbart, 2007, pp. 15-16).

In brief, individual factors such as gender, race, class, personality, learning styles, and background experiences interact to influence orientation and cognition in mathematics. Moreover, evidence suggests that temperament may be a critical individual learning factor. Equitable mathematics teaching requires the consideration of the unique effects of these individual context factors.

Environmental Context. Environmental factors interact with individual factors to influence the equitability of learning opportunities in mathematics. Controversy over the importance of environmental factors on learning lasted for decades, beginning with the publication of *Equality of Educational Opportunities*, more commonly known as *The Coleman Report* (Coleman et al., 1966). This study examined the achievement impact of differences between races on: school factors such as class size, access to chemistry, physics, and language laboratories, number of books in libraries, number of textbooks; teacher and principal characteristics such as type of college attended, years of teaching experience, salary, maternal education level, vocabulary ability, and dispositions; and student characteristics such as parental

background, presence of parents at home, size of family, parental expectations, parental involvement, and socio-economic status. *Equality* fundamentally altered definitions of equality from simply comparing resource “inputs” to analyzing the effects of inputs on educational achievements (Coleman, 1967a). Coleman (1967b) considered the complexity of implicit assumptions present within an input-based notion of equality:

It is one thing to take as given that approximately 60% of an entering high school freshman class will not attend college; but to assign a particular child to a curriculum designed for that 60% closes off for that child the opportunity to attend college. Yet to assign all children to a curriculum designed for the 40% who will attend college creates inequality for those who, at the end of high school, fall among the 60% who do not attend college... there is a wide variety of different paths that adolescents take on the completion of secondary school (Coleman, 1967b, p. 9).

Instead, *Equality* examined inequality based on five different criteria: degree of racial segregation, allocation of resources, teacher orientations, weighted resource inputs based on achievement predictability, and output (e.g., achievement, career choice) differences (Coleman, 1968). *Equality* found that student characteristics accounted for the majority of variance in achievement and of the impact of teacher characteristics on learning. For example, Coleman et al. (1966) reported that teacher variables accounted for 2.06% of the variation in mathematics achievement for Black students but only 0.61% for White students (p. 294). They concluded that “variations in school quality are not highly related to variations in achievement of pupils” (p. 297). However, technology capabilities of the time limited the researchers’ analytic capabilities (Stringfield & Teddlie, 2004). Later studies (e.g., Bryk & Raudenbush, 1988; Raudenbush &

Bryk, 1984) took advantage of technological advancements by conducting multilevel analyses on the subsets of the Coleman et al. (1966) data set:

The results were startling — 83% of the variance in [learning] growth rates was between schools. In contrast, only about 14% of the variance in initial status was between schools...this analysis identified substantial differences among schools that conventional models would not have detected (Raudenbush & Bryk, 2004, pp. 9-10).

Recent studies have continued to emphasize the importance of environmental factors on students learning. Hegedus and Kaput (2004) found that the way classroom activities are organized affects the potential depth of student understanding. Cobb, Gresalfi, and Hodge (2009) found that cultures within a classroom influence the development of personal identities in mathematics. LaRocque (2008) found that student perceptions of the classroom environment are associated with reading and mathematics achievement. She noted that the interaction of perception with gender was not statistically significant but that the interaction of perception with grade level was significant. McMahon, Wernsman, and Rose (2009) agreed with LaRocque's findings that perceptions of classroom difficulty are strong predictors of mathematics and science self-efficacy. Bong (2008) found that classroom goal structures influence student perceptions of mathematics learning. She also found that relationships influence perceptions of learning. In like manner, Carter (2008) described the impact of having a classroom climate that values the struggle of connecting mathematical concepts to current conceptions. She concluded that such a climate enhances student self efficacy and confidence. Similarly, Murayama and Elliot (2009) concluded that classroom goal structures influence the development of intrinsic motivation.

Amenkhienan and Kogan (2004) concluded that the student-teacher relationship

influences the amount of learning that occurs. Stemler, Elliott, Grigorenko, & Sternberg (2006) proposed a framework for interpersonal relationships with teachers, noting that the work of teaching is largely social in nature. Likewise, Hughes and Kwok (2007) identified teacher relationships with both parents and students as mediating factors of student motivation and achievement. They also noted an interaction between race and the amount of teacher support received. Osterman (2000) summarized research findings on the interaction of student belongingness and school and classroom conditions with motivation and achievement:

Research also tells us that conditions in the classroom and school influence students' feelings about themselves; these in turn are reflected in student engagement and achievement. Not all students experience alienation to the same extent, yet, for the most part, students and researchers describe schools as alienating institutions... While the "peer culture" may establish norms dress and behavior, it is not necessarily one that satisfies students' need for belongingness (p. 360).

Stipek (2006) added to Osterman's findings:

Learning requires effort, and one of the best predictors of students' effort and engagement in school is the relationships they have with their teachers... To promote high academic standards, teachers need to create supportive social contexts and develop positive relationships with students (p. 46).

Accordingly, the impact of the learning environment and student perceptions of that environment interact with individual context but also act as a distinct component to student learning.

Putting the Model Together

Figure 1 offers a pictorial interpretation of how the characteristics of a task and of a student's knowledge framework may operate within a mathematics classroom learning environment. Procedural knowledge isolated from conceptual knowledge and the connections between ideas results in instrumental understanding (Skemp, 1976/2006). Instrumental understanding may result in a cycle of misconceptions and faulty reasoning reinforcing each other and weakening a student's knowledge framework for understanding future tasks. When conceptual knowledge and procedural knowledge develop together, they reinforce each other and strengthen a student's knowledge framework for future tasks. When students complete a task, teachers have limited opportunities to assess the knowledge framework and thought processes that lead to a response; instead, assessment usually focuses on whether or not a response was correct. Unfortunately, correct responses do not necessarily indicate that a student understands the mathematical concepts completely. Figure 10 therefore includes the possibility that correct responses can be produced even with erroneous reasoning, and if unchecked, that reasoning will reinforce misconceptions and erroneous reasoning, thereby weakening a student's knowledge framework for understanding future tasks.

Resnick (1983) suggested that errors often occur when students look for meaning in situations where the given information is incomplete. In such cases, students often attempt to use their prior knowledge to fill in the gaps and yielding misconceptions. Thus Resnick (1983) found that working through these difficulties may be a normal part of the learning process and that combating misconceptions and faulty reasoning must become an expected part of the struggle that is so critical to deep conceptual learning as Hiebert and Grouws (2007) later pointed out. Moschkovich (1998) agreed with Resnick when she noted refinement of understanding as a

primary goal of teaching: “We need to understand the process of conceptual change that enables learners to transform and refine their conceptions to more closely fit with the desired understanding” (p. 209).

Tracing the root causes of errors and recognizing erroneous reasoning requires an examination of student explanations about their reasoning processes. Previous interventions targeting specific errors or the underlying reasoning have met with limited success (e.g., Clement, 1982; Fisher, 1988; Phillippe, 1992; Rosnick & Clement, 1979), possibly because these interventions may have targeted the error instead of the latent reasoning and misconception that led to the error. Furthermore, students with misconceptions and faulty reasoning may produce correct answers; as a result, interventions focusing on errors may miss unobservable erroneous reasoning.

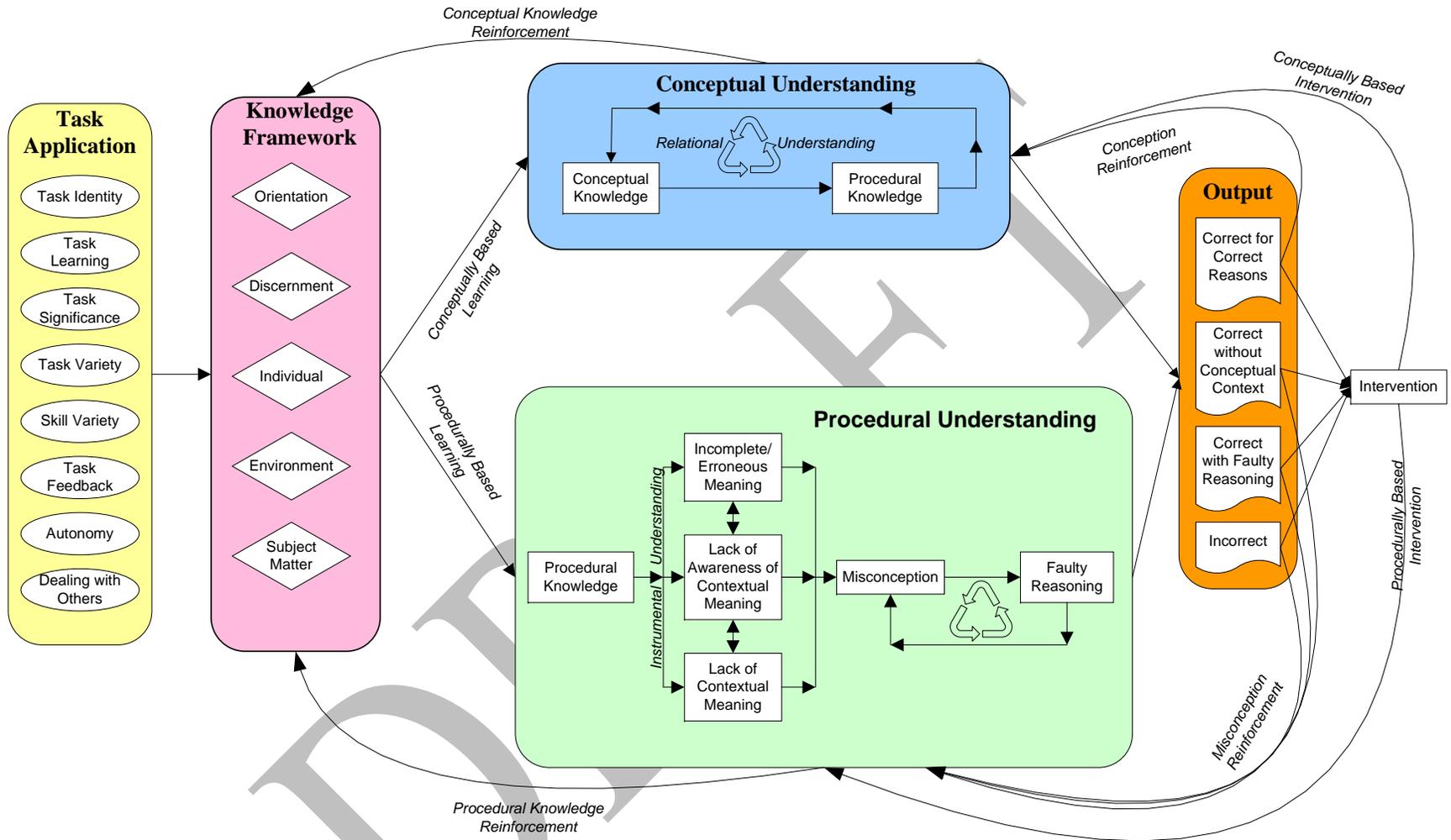


Figure 1. Pathways of Conceptual Learning in Mathematics.

Methods

The measurement of mathematical misconceptions is inherently problematic due to the latent nature of those misconceptions. For example, Zawojewski and Shaughnessy (2000) pointed out the inadequacy of simple multiple choice tests to identify the thought patterns that result in a particular answer. Instead, they recommended including a qualitative component to each question to provide clues to underlying student thinking. In order to include that strategy in the instruments used in this study, an initial assessment of student responses was necessary to determine the source of student reasoning errors. For example, were reasoning errors occurring on a particular due to a lack of relational understanding, despite having relational understanding, or due to a more fundamental misunderstanding of foundational mathematical ideas? The results of that analysis were used to code errors for the subsequent quantitative analyses. As such, the design of this study falls within the mixed methodology design as described by Tashakkori and Teddlie (1998).

Research Design

The present study employed a randomized pre-posttest design with a control group (Equation 1) with 19 mathematics teachers recruited from four schools in three Kentucky school districts with 1,142 students enrolled in their 53 algebra and geometry classes.

$$\frac{R \quad O_{ATMI/MAI} \quad O_{NAEP} \quad X \quad O_{ATMI/MAI} \quad O_{NAEP}}{R \quad O_{ATMI/MAI} \quad O_{NAEP} \quad O_{ATMI/MAI} \quad O_{NAEP}} \quad (1)$$

The research design (i.e., random assignment and use of a pretest as a covariate) addressed potential selection and maturation threats to internal validity. The outcome of interest was the decline of misconception errors over the treatment period.

Instrumentation

Three instruments were used to measure student mathematics knowledge, student attitudes toward mathematics, and student metacognitive knowledge and skills. The mathematics knowledge instrument was used to account for pre-existing mathematics knowledge and ability. It was also used to analyze error response patterns to determine which errors emerged from mathematical misconceptions or from non-conceptual reasoning errors.

Mathematics Misconceptions Instrument. Items for the mathematics misconceptions instrument were gathered from National Assessment of Educational Progress (NAEP) released items (U.S. Department of Education, 1996, 2005, 2007). Although all 17 items remained as given by NAEP, a prompt was included for each question asking students to explain how or why they chose their response.

NAEP items are rigorously developed, using review boards, pilot testing, classical test theory, and Item Response Theory to analyze item performance (U.S. Department of Education, 2008a). These items were deemed to have high content validity for the NAEP-associated content areas.

These items included rational number, probability, algebra, and geometry content. Table 8 provides a description of each NAEP item used in the assessment instrument along with the reported reliability coefficients for each item block (U.S. Department of Education, 2008b, 2008c, 2008d, 2008e). These items were chosen based on two criteria: (1) The item content matched the foundational concepts that research has suggested connect rational number, probability, algebra, and geometry misconceptions closely enough to be able to detect intervention effects; and, (2) The item content and wording did not so closely match the activities and problems in the probability unit that the

treatment group would receive an unfair advantage over the control group. Table 2 presents the classical test theory difficulty coefficient (i.e., percent correct), the NAEP classification of difficulty and complexity level of each item with respect to the intended grade level of the item, and the internal consistency of the associated block of items as they appeared on the NAEP instruments.

DRAFT

Table 2
Reported NAEP Item Performance

Item	Release Year	Content Strand	Content	Percent Correct	Grade Level	Difficulty	Complexity	NAEP Block	Cronbach Coefficient α
1	2007	Probability	Relative versus absolute comparison	45%	Grade 4	Medium	Low	M7	0.80
2	2005	Probability	Determine Conditional Probability	49.5%	Grade 12	Medium	Low	M12	0.73
3	2007	Probability	Repeated Sampling Probability	60%	Grade 8	Medium	Low	M11	0.76
4	2005	Probability	Dependent probability	18%	Grade 8	Hard	Moderate	M12	0.75
5	2007	Algebra	Convert temperature units	35%	Grade 8	Hard	Low	M9	0.80
6	2005	Algebra	Effect of variable change	34%	Grade 8	Hard	Moderate	M3	0.76
7	1996	Algebra	Additive versus Multiplicative Structure	58%	Grade 8	Medium	-	M3	0.53
8	2007	Algebra	Solve algebraic word problem	47%	Grade 8	Medium	Moderate	M11	0.76
9	2007	Geometry	Determine if a shape is a parallelogram	26%	Grade 8	Hard	Moderate	M11	0.76
10	2005	Geometry	Area of shaded figure	77%	Grade 8	Easy	Low	M4	0.77
11	2005	Geometry	Find dimensions from scale drawing	85%	Grade 12	Easy	Moderate	M12	0.75
12	2005	Rational Number	Rational Number Quantity Meaning	66%	Grade 12	Easy	Low	M3	0.73
13	2005	Rational Number	Given the scale, determine length of side	56%	Grade 12	Medium	Low	M4	0.79
14	2007	Rational Number	Arrange fractions in ascending order	49%	Grade 8	Medium	Low	M9	0.80
15	2007	Rational Number	Determine fraction of figure shaded	89%	Grade 8	Easy	Low	M11	0.76
16	2007	Algebra	Determine equation to represent table.	54%	Grade 8	Medium	Moderate	M7	0.78
17	2005	Probability	Determine amount from probability	40%	Grade 8	Medium	Low	M4	0.77

Since the items were chosen from different blocks, the NAEP-reported coefficients do not necessarily represent the internal consistency of the new instrument compiled for the present study. Therefore, the pooled internal consistency of the new instrument was re-assessed using the pretest data ($\alpha = 0.791$, 95% CI [0.773, 0.808]) and the posttest data ($\alpha = 0.772$, 95% CI [0.751, 0.773]) and found to have adequate internal consistency. The correlation of each item (Table 9) between the pre- and post-tests were computed to measure test-retest reliability (i.e., stability). The correlations were moderate and significant ($p < 0.001$) for all items except Item 17, which was only significant at the 93% confidence level ($p = 0.068$). Overall, the stability of the items appeared to be acceptable (Table 3).

Table 3
Stability Correlations Between the Pre- and Post-Test Data for each Item

Item	1	2	3	4	5	6	7	8	9	10	11
Correlation	0.491	0.173	0.273	0.310	0.277	0.217	0.422	0.308	0.279	0.354	0.385
Item	10	11	12	13	14	15	16	17			
Correlation	0.354	0.385	0.268	0.154	0.460	0.300	0.325	0.083			

Content validity of content area alignment to national, state, and local standards was evaluated by the NAEP Validity Studies Panel. Daro, Stancavage, Ortega, DeStefano, & Linn (2007) examined the content coverage, skill coverage, alignment to NAEP framework, lack of philosophical bias, lack of ability bias, and representativeness of information provided about students. They found that 96% of NAEP 2005 and 2007 items demonstrated adequate or marginal quality.

Item Response Theory (IRT) was applied to measure the characteristics of difficulty, discrimination (i.e., the ability to distinguish between groups, in this case, ability levels), and guessing for each item. IRT, unlike Classical Test Theory (CTT), focuses on the correctness or incorrectness of each item individually rather than a raw cumulative score (Baker & Kim, 2004). In CTT, difficulty is defined as the percentage of correct responses for an item. CTT

discrimination is typically measured as the point-biserial correlation for each item. One problem with CTT is the circular dependence of observed scores and samples (Fan, 1998).

IRT is based on the item characteristic curve, which is computed using a logistic function. The curve can be computed as a Rasch model (1 parameter, item difficulty), 2PL (2 parameters, item difficulty and discrimination), or 3PL (3 parameters, item difficulty, discrimination, and guessing). The logistic function for the 3-PL curve is

$$P(\theta) = c + (1 - c) \frac{1}{1 + e^{-a(\theta - b)}} \quad (2)$$

where:

a represents the discrimination coefficient

b represents the difficulty coefficient

c represents the guessing coefficient

θ represents the ability level of the respondent

ParScale 4.1 (Muraki & Bock, 2002) uses an iterative process to compute the item characteristic curve. In the first iteration, ability levels (θ) for each subject were computed. These values become the starting point for the second iteration, which is used to compute the values for a , b , and c . The guessing coefficient, c , was estimated as $c = 0$ for all 17 items. Therefore, the model reduced to a 2PL curve, and the values for a and b were computed for each item (Table 4).

Table 4
IRT Coefficients for NAEP Items

Item	Discrimination, a	Difficulty, b
1	0.570	-0.618
2	0.630	0.379
3	0.693	-0.205
4	0.962	-0.090
5	0.827	-0.175
6	0.570	0.930
7	1.192	-0.273
8	0.730	0.430
9	0.428	-1.308
10	0.918	-0.446
11	1.224	-0.080
12	0.879	0.149
13	0.876	0.246
14	0.861	0.311
15	0.883	-0.554
16	1.036	0.203
17	0.744	0.525

Note: SE for all 17 items for both a and b was < 0.001

The item characteristic curves (Figure 2) can be used to compare the behavior of each item. For each curve, the horizontal axis represents difficulty, b , and the vertical axis represents the ability, θ .

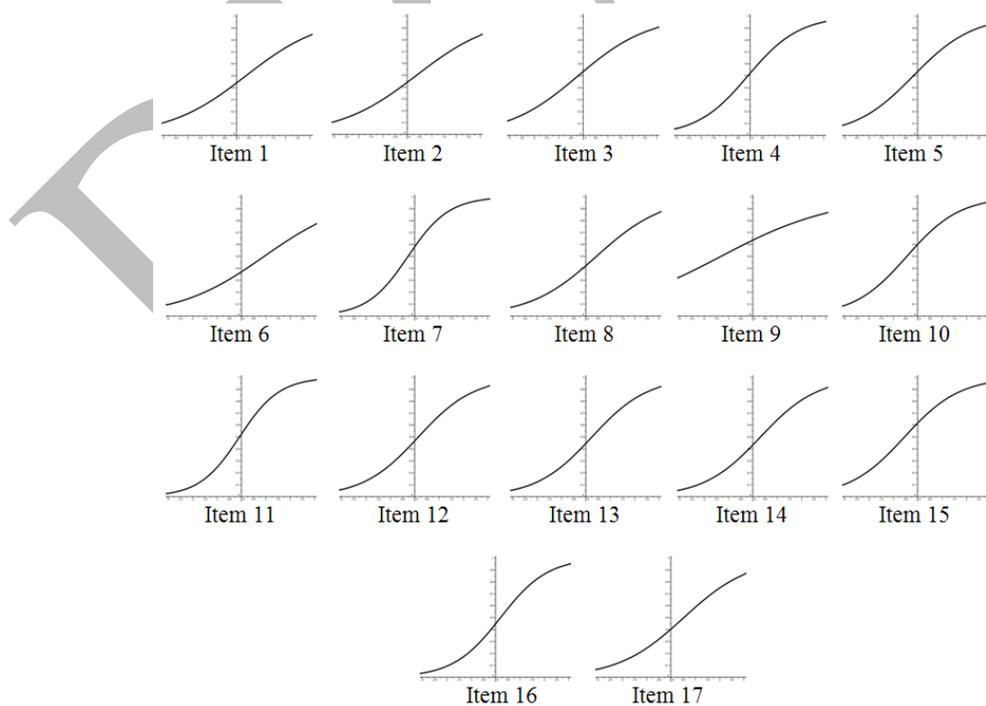


Figure 2. Item Characteristic Curves for NAEP Mathematics Knowledge Instrument

None of the curves in Figure 2 leveled off at the top or bottom, indicating no ceiling or floor effects. The discrimination of items on the mathematics knowledge test ranged from 0.428 (less differences between high and low ability students) to 1.192 (more differences between high and low ability students). The difficulty ranged from -1.308 (Easy) to 0.930 (Hard). Because the discrimination levels did not approach 0 (no differences between high and low ability students), and the difficulty levels did not indicate that any items were extremely easy (approaching -3 and +3), the item characteristics were considered appropriate for the planned analyses.

Mathematics Attitudes Inventory. Student orientation was measured using the Attitude Toward Mathematics Inventory (ATMI; Tapia & Marsh, 2004). This instrument was selected because its subscales have been extensively analyzed to establish high reliability and content validity (Tapia & Marsh, 2004). The subscales for this inventory were developed from multiple literature sources to maximize concurrent construct validity. According to the Tapia and Marsh report, the ATMI measures four orientation constructs. Factor 1, self confidence, consists of 15 items with a reported Cronbach alpha of 0.95. Factors 2 and 3, perceptions of the value of mathematics and enjoyment of mathematics, each contain 10 items with a Cronbach alpha of 0.89. Factor 4, motivation to learn mathematics, contains five items with a Cronbach alpha of 0.88.

The internal consistency for the full instrument and each subscale was measured using the present study data to determine their reliabilities (Table 5).

Table 5
Internal Consistency Reliability for ATMI

Scale	Observed Cronbach Alpha	95% Confidence Interval
Full Instrument	0.943	[0.938, 0.949]
Factor 1: Self Confidence	0.909	[0.901, 0.918]
Factor 2: Value	0.876	[0.864, 0.888]
Factor 3: Enjoyment	0.798	[0.778, 0.817]
Factor 4: Motivation	0.824	[0.806, 0.842]

The observed reliability coefficients for the present study data appeared to be comparable to those reported by Tapia and Marsh (2004) and had values higher than the typical threshold of 0.7 (Urbina, 2004). The reliabilities were, therefore, determined to be acceptable.

MetaCognition Inventory. Student metacognition knowledge and skills were measured using the Metacognitive Awareness Inventory (MAI; Appendix P; Schraw & Dennison, 1994). This instrument was selected because of its unique subscales of metacognition, knowledge of cognition and regulation of cognition and because it has been rigorously tested through two experiments to establish concurrent construct validity for each block of items. Three types of knowledge are measured as components of knowledge of cognition: (1) Declarative knowledge, defined as knowledge of learning and of one's own cognitive skills and abilities; (2) Procedural knowledge, knowledge of how to use various cognitive strategies; and, (3) Conditional knowledge, knowledge of when to use particular cognitive strategies and why those strategies should be used. Under the regulation of cognition, five components are measured: (1) Planning, including goal setting and allocation of resources; (2) Organizing and managing information; (3) Monitoring, reflection on cognitive processes during a learning task; (4) Debugging, strategies for correcting performance errors or assumptions; and, (5) Evaluation, reflection on cognitive processes after a learning task is completed (Schraw & Dennison, 1994). Schraw and Dennison also reported high internal consistency for the whole instrument ($\alpha = 0.93$) and both

metacognition factors ($\alpha = 0.88$). The internal consistency for the full instrument and each subscale was measured using the present study data to determine their reliabilities (Table 6).

Table 6
Internal Consistency Reliability for MAI

Scale	Observed Cronbach Alpha	95% Confidence Interval
Full Instrument	0.946	[0.941, 0.951]
Factor 1: Knowledge of Cognition	0.870	[0.857, 0.882]
Declarative Knowledge	0.744	[0.718, 0.768]
Procedural Knowledge	0.615	[0.573, 0.654]
Conditional Knowledge	0.668	[0.633, 0.700]
Factor 2: Regulation of Cognition	0.924	[0.917, 0.932]
Planning	0.727	[0.699, 0.752]
Organizing	0.788	[0.767, 0.808]
Monitoring	0.735	[0.708, 0.760]
Debugging	0.694	[0.662, 0.725]
Evaluation	0.673	[0.639, 0.705]

The observed reliability coefficients for the full instrument and two main factors demonstrated high internal consistency. Several of the sub-factors showed marginal reliabilities ($\alpha < 0.7$). These reliabilities were similar to those reported by Schraw and Dennison, who only reported the two main factors as reliable. Therefore, only the two main factors were used in the subsequent analysis of contextual factors.

Missing Data

Rubin (1987) classified missing data due to non-response as either unit non-response, meaning that the subject refused to answer any of the items, and item non-response, meaning that the subject skipped questions. The ATMI and MAI surveys of the present study included both types of non-response. Forty six students did not respond to any items on either survey; 63 additional students did not respond to a majority of items, ending at various points throughout the survey (Table 7). The format of the questionnaire may shed light on the most typical pattern of unit non-response: The front of the survey form included the ATMI and the first three

questions of the MAI. Questions 4 – 52 of the MAI (i.e., the back of the survey) were the most commonly skipped questions. Based on this pattern, which may well be the result of bias in the non-response patterns, I concluded that most non-response on the MAI was due to the presentation format of the instrument.

The NAEP achievement data, both pre- and post-test, consisted of very low proportions of missing data. On the pretest, missing data accounted for 5.6% of the entries across all items and subjects, and only 11 students (< 1%) did not respond to any items (Table 7). On the posttest, missing data accounted for 6.5% of the entries across all items and subjects, and only 12 students (1.1%) did not respond to any items.

Table 7
Sources of Missing Data

Instrument	N	Unit Non-Response (No items answered)	Item Non-Response
ATMI	964	46 (4.6%)	<ul style="list-style-type: none"> • 0 of 40 items with full data • 256 (26.6%) cases missing at least one value • 2,457 of 16,388 (6.4%) values missing
MAI	964	109 (11.3%)	<ul style="list-style-type: none"> • 0 of 52 items with full data • 316 (32.8%) cases missing at least one value • 6,060 of 16,388 (12.1%) values missing
Pretest	1142	11 (0.96%)	<ul style="list-style-type: none"> • 0 of 17 items with full data • 242 (21.4%) cases missing at least one value • 1093 of 19,278 (5.6%) values missing
Posttest	1021	12 (1.1%)	<ul style="list-style-type: none"> • 0 of 17 items with full data • 248 (24.3%) cases missing at least one value • 1121 of 17,357 (6.5%) values missing

Hierarchical linear modeling was used to examine the impact of item, student, and class characteristics on misconception errors.

The impact of knowledge filters (Figure 3) on misconception emergence will be assessed using hierarchical generalized linear models (HGLM). The Bernoulli HGLM model best fits the dichotomous nature of the outcome data (Raudenbush & Bryk, 2002). In a Bernoulli model, the

outcome variable is transformed to a logit, η , of the odds ratio, ϕ , for the outcome variable. The computed logit becomes the linear outcome variable. In this analysis, NAEP items (Level 1) are nested within students (Level 2), and students are nested within classes (Level 3).

Qualitative Analysis

The qualitative analysis served two critical functions in the present study. First, classroom observations and teacher interviews (structured around topics relating to implementation of the intervention lessons) before, during, and after the treatment periods provided data on fidelity of treatment. While observing classrooms, the researcher attempted to minimize distractions inherent to having a visitor in a classroom. In some classrooms, this goal was best met by slipping in quietly and sitting in the back of the class. In other classrooms, teachers preferred to introduce the researcher and involve him in the lesson.

Fidelity to the probability intervention lessons varied widely between teachers. Shaughnessy and Bergman (1993) pointed out that many teachers are uncomfortable with probability content; varying responses to such discomfort were expected. Some teachers preferred to revert to normal, procedural methods of teaching. In this case, the ability of a probability unit to counter misconceptions may have been reduced. Other teachers followed the lessons provided by the researcher with varying degrees of success. One role of the researcher during the treatment period was to provide assistance to the teachers throughout the intervention lessons.

The second major function of the qualitative analysis was to provide an analysis of student responses to the open response items on the mathematics knowledge assessment. This analysis was used to assess hypotheses of previously identified misconception patterns and advance the understanding of the relationship between mathematical misconceptions and

reasoning errors. This analysis was conducted from the constructivist point of view using a narrative analysis (Creswell, 2007; Patton, 2002) of symbolic interactions (i.e., the symbols used to provide meaning to students), semiotics (i.e., how signs and symbols are used to convey meaning), and hermeneutics (i.e., how students interpret signs and symbols).

These qualitative analyses were fundamental for establishing the context for all subsequent quantitative analysis. The first analysis provided evidence of treatment fidelity and intervention effectiveness. The second analysis provided a foundation for the interpretation of all quantitative findings.

Quantitative Analysis

The impact of item, student, and class characteristics on misconception errors was investigated using hierarchical modeling. The Bernoulli HGLM model best fit the dichotomous nature of the outcome data (Raudenbush & Bryk, 2002). In a Bernoulli model, the outcome variable is transformed to a logit, η , of the odds ratio, ϕ , for the outcome variable. The computed logit becomes the linear outcome variable. In this analysis, NAEP items (Level 1) were nested within students (Level 2), and students were nested within classes (Level 3).

Scale variables were centered to facilitate interpretation of the intercepts and slopes. In the student Level 2 equations, centering occurred at the group level (noted by the subscripts jk and $\bullet k$), causing the intercepts and slopes to be interpretable as student deviation from the classroom average (i.e., intercept represents the average student in the same class on all predictor variables; slope for any variable represents the impact of being an above or below average student in the class). In the classroom Level 3 equations, scale variables were grand mean centered, meaning that the overall mean is subtracted from each classroom mean (noted by the subscripts k and \bullet). Grand mean centering changes the interpretation of the classroom Level 3

intercepts and slopes just as group centering did on the student Level 2 intercepts and slopes. At the classroom level, intercepts now represented the value for the average classroom on all predictor variables, and slopes represented the impact of being in an above or below average classroom. The regression coefficient for each classroom variable therefore represented the effect of a variable on the impact of each corresponding student variable.

Results

Identifying Misconception Patterns

A sub-sample was chosen for a qualitative analysis of patterns of misconception responses on the NAEP-based mathematics knowledge test. On all items, students were asked to provide an explanation for their response choice. Approximately 74% of the sample left these explanations blank. To improve representativeness of the overall sample, the qualitative sub-sample was chosen using purposive stratification across classes; specifically, tests were chosen to be part of the sub-sample if they filled in the explanation section of the test for most items. Such a sampling technique produced a selection bias — students who completed their explanations were more likely to choose the correct answer, resulting in a reduced sample for each distractor to each item. To help manage this bias, the sampling procedure continued until all distractors for each item were represented ($N = 72$). Division of items by content area (i.e., algebra, geometry, probability, and rational number) transferred directly from the NAEP classification of each item.

The following descriptions focus primarily on student explanations of errors; however, the analysis began with a recognition that correct responses do not necessarily indicate conceptual understanding. For all items in which the correct response explanations are not discussed explicitly, the explanations by students indicated that they did, in fact, understand the

concept not understood by students who chose incorrect responses. The thick description provided in this analysis was used to establish trustworthiness for the coding of misconception responses.

Misconceptions on Algebra Content Knowledge Items

Algebra items (i.e., Items 5, 6, 7, 8, and 16; Appendix N) included distractors that reflected misconceptions about additive/multiplicative structures, the meaning and interpretation of variables, and the meaning of rational numbers.

Item 5 response patterns. Item 5 described the formula to convert temperature from Fahrenheit to Celsius in words and then asked students to convert 393°F to Celsius. I hypothesized that choices A, B, and C would represent misconceptions about the meaning of rational numbers. Student responses confirmed this hypothesis. For example, one student chose A, “Because $393 - 32 = 361$; $5/9 = .55$, so you divide 361 by $5/9$, answer is 656.3, to the nearest degree is 650.” This explanation represents the explanations of others who chose A, indicating that students who chose A did so because they divided by the rational number rather than multiplying, not realizing that the resultant rational number, $5/9$ of 361, should be smaller than 361.

I hypothesized that Choice B for Item 5 would result from the ignoring of the denominator, and again, student explanations confirmed this hypothesis. The most explicit case of this type chose B, “Because $393 \times 5 = 1805$.” This student also failed to note that he/she had subtracted the 32 from the 393 properly; 361×5 is 1805 while 393×5 is actually 1965. I decided that this particular error was simply one of reporting rather than a misconception, so it was excluded for the purposes of this analysis. Another student who chose B stated, “ $361 \cdot 5/9 = 200$ $5/9 = 200 \times 9 = 1800 + 5 = 1805$.” This student failed to realize that the rational number was

accounted for by the 200 and continued to try to incorporate the fraction, ultimately doing so by misusing both numbers. From this question, I considered how the student had correctly computed the 200 if he/she did not understand how to use the $\frac{5}{9}$ later. My best guess was that the student used a calculator for the first computation, but thought that 200×9 would be an easy calculation, so he/she did the last steps by hand and did not check them on the calculator. Although this conclusion is wholly speculative, if true, it may suggest that the use of calculators to explore the meaning of rational numbers may open an avenue for addressing student conceptions and perceptions of rational numbers.

I also hypothesized that Choice C would represent a misconception about rational numbers, specifically, that students would choose C by ignoring the rational number altogether. Student responses also verified this hypothesis. Students who chose C justified their response with statements such as, "Divide 393 and 32."

Originally, I hypothesized that E would not represent a similar misconception as those for A, B, and C on Item 5. Student responses, however, contradicted this hypothesis. Students who chose E also ignored the denominator and misused the numerator as did students who chose B. For example, students justified choice E with statements such as, "Divide the numbers," specifically 361 by 5. Therefore, Choice E was added to the misconception choices for Item 5.

These interpretations of rational numbers appeared to typify student responses to rational numbers. I generalized these patterns of rational number interpretation into five types:

1. Rational number is understood to be a single quantity, but confusion about the meaning of that quantity results in the application of the wrong operation or the correct operation(s) to the wrong quantities (e.g., Divide instead of multiply, multiply by the wrong number). This error connects to rational

- number meaning misconceptions identified by Fosnot and Dolk (2002).
2. Reverse the role of the numerator and/or denominator. This error is similar to those described by Baturu (1994), Behr et al. (1992), and Lamon (1999).
 3. Ignore either the numerator or denominator (as in Green, 1983b; Watson & Shaughnessy, 2004).
 4. Ignore the numerator/denominator AND reverse the role of the remaining part of the rational number (e.g., Divide by the numerator)
 5. Ignore the rational number altogether. This error appeared to connect to variable number misconception described by Küchemann (1978) in which students ignored the presence of variables.

While this categorization of rational number meaning errors may not account for every rational number meaning error for every problem, it may serve as a foundation for exploring rational number meaning errors in other problems/contexts. Additionally, this list appears to be hierarchical; that is, a Type 2 rational number misinterpretation may represent a greater degree of confusion about the meaning of rational numbers than Type 1, as would Types 3, 4, and 5 over a Type 2.

Item 6 response patterns. Item 6 offered students a linear function ($y = 4x$) and asked about the change in y based on an increase of two units to x . Prior to the present analysis, only Choices D and E were hypothesized as misconception responses — I expected students with additive/multiplicative structure misconceptions to choose D by squaring the independent variable coefficient and E by doubling that same coefficient (as described by Warren, 2000). Explanations from students who chose these two responses supported this expectation. For example, students who chose E typically showed their calculation as “ $4 \times 2 = 8$.” Alternatively,

students who chose A, the correct answer, also performed this same calculation but knew to add it to the overall y value rather than making it a new coefficient. Therefore, D and E remained misconception responses in the coding procedures. Additionally, explanations for choices B and C also indicated misconceptions in student reasoning about additive and multiplicative structures. For example, students who chose B stated, “ $4 + 2 = 6$ ” or “It also increases by 2,” similar to patterns found by Warren (2000) and Moss et al. (2008). Likewise, students who chose C offered one of three justifications, all of which represented a misconception about how to handle an additive structure in an algebraic equation. The first type of explanation demonstrated a reliance on the balance-beam principle of algebraic equations, stating something such as, “You basically add 2 more to the other side,” “Because both sides need to be the same, or “Because if one increases, so does the other.” The second type of explanation showed that some students chose C because they thought that the change of two should be added to the coefficient, stating rationales such as, “Because $4 + 2 = 6$ that is 2 more than the original amount” or “Because $4x + 2 = 6x$.” The third type of response to support Choice C indicated that students knew that the change in y should be additive, but they failed to understand the role of the coefficient in that change. These students justified their choice of C by asserting, “It would be $y = 4x + 2$.” As a result of this analysis, the coding of misconception responses was expanded to include choices B and C. These errors appeared to occur because students were relying on procedures isolated from meaning and connections between ideas. The framework in Figure 10 may shed light on how these errors emerged. For example, students who relied on the balance beam principle for solving algebra equations did not seem to understand why such an approach works and what it means to a particular context. These students appeared to demonstrate procedural knowledge with instrumental understanding (Skemp, 1976/2006). As a result, these students developed algebraic

misconceptions about how, when, and why the algebra balance beam works. These misconceptions about the balance beam led to faulty reasoning that may have reinforced the balance beam misconceptions.

Item 7 response patterns. Item 7 asked students to choose an expression to represent the situation, “A plumber charges \$48 for each hour and an additional \$9 for travel.” The correct response, choice E, uses the \$48 per-hour charge as the coefficient to the number of hours and adds the \$9 travel fee as a one-time charge. Every distractor response was hypothesized to represent a misconception about additive/multiplicative structures. Student explanations verified this prediction. For example, students who chose A interpreted each charge as an “additional” charge. Likewise, students who chose B thought the calculation resulting from such an expression should be “added on to the original.” They also believed that both charges should be multiplied by the number of hours. Students who chose C knew that a charge should be added and another multiplied, but they reversed the quantities. Finally, students who chose D understood that the expression should represent “48 times the hours plus 9,” but they did not understand how to translate those words into an expression. I concluded that choice D could possibly represent variable misconceptions as well as additive/multiplicative structure misconceptions, but since the analyses of the present study will not differentiate between types of misconceptions, D was left as a misconception response.

Item 8 response patterns. Item 8 presented students with the following scenario: “Carmen sold 3 times as many hot dogs as Shawn. The two of them sold 152 hot dogs altogether. How many hot dogs did Carmen sell?” Originally, I hypothesized that choice B would represent the reversal error, reflecting the wrong person’s amount, similar to the reversal error identified by Clement (1982). The present analysis revealed that such was not the case in this sample. In fact,

students who chose B provided correct equations such as “I did $3s + s = 152$ and add the s to the $3s$ to get $4s$ and divided 152 and 4 by 4 and got $s = 38$.” Not one student who chose B in the sub-sample indicated that they thought the 38 represented Carmen’s amount. As a result, I concluded that choosing B did not represent a variable meaning misconception so much as a careless error; specifically, not catching that the question asked for Carmen’s amount instead of Shawn’s.

Almost twice as many students in the sub-sample chose C instead of B, and these students did indicate one of Küchemann’s (1978) variable interpretation errors — they unanimously ignored the variables altogether and simply divided 152 by 3 . Students who chose E also ignored the variables entirely, showing a similar calculation to that of C. For example, students who chose C justified their answers with statements such as, “Because $152 \div 3$ gives you 50.8 , round and you get 51 .” Likewise, students who chose E wrote statements such as, “Because $50 + 50 + 50 = 150 - 2$.” In both cases, students failed to recognize the role of the variable in partitioning the total amount. Other students who chose C and E demonstrated Küchemann’s Level 1 interpretation, evaluating the variable using trial and error. Therefore, I concluded that choices C and E should represent the variable misconception rather than the original choice B.

Item 16 response patterns. Item 16 presented students with a table of values and asked them to determine the function that best modeled the data. I hypothesized that choice A would represent a misconception about the nature of the functional relationship; student responses verified this expectation. Students who chose A made statements that indicated an understanding of a relationship, but they looked at the relationship backwards, i.e., they thought of n as the dependent variable rather than the independent variable, similar to the reversal error in Clement (1982). Although such an inversion might also be due to a rational number meaning misconception (i.e., doubling rather than halving, as in Item 5, error 2, Baturo, 1994; Behr et al.,

1992; Lamon, 1999), the purpose of the present analysis is not to distinguish between misconception types but rather to reflect the presence of misconceptions in a content area. Therefore, choice A remained a misconception response in the analysis. This overlap indicates the possibility of an underlying multicollinearity across content area misconceptions resulting from the influence of the underlying misconceptions.

Student explanations for choosing C or D on Item 16 also indicated the presence of variable misconceptions. Students who chose C explained that they had only used the first column of values to determine the equation, concluding that, “You subtract by the n and that equals p .” For choice D, students explained that they thought the number of days was one of the variable quantities of importance, “Because the days matter too.” These students then used the first column of values to conclude that subtraction held the key to solving this problem. As a result of this analysis, choices C and D were included as misconception responses for Item 16.

Misconceptions on Geometry Content Knowledge Items

The geometry items on the NAEP instrument (i.e., Items 9, 10, and 11; Appendix N) examined student misconceptions about spatial reasoning and the meaning of rational numbers.

Item 9 response patterns. Item 9 presented students with a rectangle and asked them whether the figure should be classified as a parallelogram. Students who chose the correct response, A, did so because, “It has parallel sides” or “It has equal sides.” This explanation indicated that these students were operating at least at Van Hiele Level 1, in which students recognize that figures have characteristics and properties. Students who chose B, on the other hand, indicated operating at Van Hiele Level 0, in which students rely on visual recognition of shapes. For example, students made statements such as, “Parallelograms are crooked ” and “The figure she drew has right angles.” Other students indicated that they thought that being

a rectangle and square excludes a shape from being a parallelogram through statements such as, “It is a square” and “No. It is a rectangle.” These types of errors indicate a fundamental spatial reasoning misconception resulting from low Van Hiele levels of understanding (Clements & Battista, 1992; Crowley, 1987). Based on this evidence, I retained choice B as a misconception response.

Item 10 response patterns. Item 10 presented students with a shaded figure within a grid of centimeter squares. I hypothesized that Choice D would represent a spatial reasoning misconception, in which students would rely on the lengths to compute the area rather than on the meaning of area, similar to Clements and Battista’s (1992) 9th and 11th most common spatial reasoning misconceptions. Explanations by students who chose D verified this hypothesis with statements such as, “Because you multiply the length and width of squares.” Additionally, to arrive at Choice D, students also needed to be confused about the length of the diagonals; in this case, they evidently chose to add them as a little more than 1, then rounded the length to an even 7 cm. They also failed to recognize that the width on one end was 2 cm while on the other end it was 3 cm. Therefore, this choice was retained as a misconception choice.

Additionally, an analysis of Choice C explanations revealed the presence of additive/multiplicative structure and spatial reasoning misconceptions, or Van Dooren et al.’s (2003) illusion of linearity applied to geometric shapes. Students who chose C explained that they simply “counted them all up,” approximating the diagonal lengths as 1.5 cm and computing a perimeter rather than an area. Therefore, I added Choice C as a misconception choice for Item 10.

Students who chose A for their response to Item 10 also revealed an error in spatial reasoning. These students recognized that they needed to count shaded areas, but they counted

only the wholly shaded squares and ignored the half shaded squares, as indicated by statements such as, “There are 9 squares that are fully shaded” and “The image is mainly 3 by 3 so 9 sq centimeters, $a = 3 \bullet 3 = 9$.” These explanations led me to believe that Choice A resulted from faulty reasoning, but not necessarily a misconception — they knew to add areas, and they appeared to recognize that area is the space inside a closed figure. Therefore, Choice A was not added to the misconception choices for Item 10.

Item 11 response patterns. For Item 11, students were asked to convert the dimensions of an object from one unit of measure to another. I hypothesized that the difficulty in Item 11 for students lay in the recognition that the dimensions represented two quantities rather than one, so the misconception of interest was the way students ignored these variable quantities. Some students ignored one of the variables; others thought that the variables were simply labels and therefore traded labels in a one-to-one relationship. Prior to this investigation, I hypothesized only that choice A would represent a misconception response, indicating the use of the variable as a label. Explanations by students who chose A confirmed this hypothesis, indicating that the values “don’t change.”

Additionally, students who chose B and C also indicated that only one variable needed to be accounted for in the computation. Students justified their choices with explanations such as, “I did $5 \times 3 = 15$, then left 3 the same.” Therefore, choices B and C were added to the list of misconception responses for Item 11.

From this investigation, misconception responses for geometry items were expanded to include additional choices. Because misconceptions such as additive/multiplicative structure were also evident in algebra items, items from both content domains may co-vary in the SEM analysis.

Misconceptions on Probability Content Knowledge Items

Probability items (i.e., Items 1, 2, 3, 4, and 17; Appendix N) on the NAEP instrument examined student errors in probability rooted in absolute/relative comparison, additive/multiplicative structure, spatial reasoning, and rational number meaning misconceptions.

Item 1 response patterns. For Item 1, students were asked to determine which picture represented the greatest probability. I hypothesized that misconception responses would follow the patterns identified by Shaughnessy & Bergman (1993) and Watson & Shaughnessy (2004): Students with probability misconceptions would focus on the number of black marbles rather than the ratio of black to white marbles, or, confusing absolute and relative comparisons. Students demonstrated this misconception precisely as expected. For example, students who chose answer C stated that its dish “contains more blacks” or “has the most black marbles.” Likewise, students who chose E also indicated that the number of black marbles was the only number of importance. Students who chose “B” made two absolute comparisons, looking for a combination of the “most white and black.”

Item 2 response patterns. Item 2 presented students with a two-way table (gender by color of puppies) and asked students to compute the conditional probability of a puppy being male given that it is brown. I hypothesized that Choices B, C, D, and E would indicate a misconception about the meaning of rational numbers, as described by Bar-Hillel and Falk (1982) and Falk (1992). Student explanations for these responses verified this hypothesis.

Students who chose B on Item 2 ignored the condition of being brown altogether and instead gave the probability of being male, providing explanations such as “Because it’s seven puppies and two of them are male” or “2 total males and 7 total puppies; chance a male will be

picked $2/7$.” These students appeared to be unsure of how to incorporate the brown condition into the probability quantity.

Students who chose C and E on Item 2 gave explanations that indicated confusion between part-part relationships (i.e., odds) and part-whole relationships (i.e., probability). These students justified their choices with statements such as, “There’s 1 male and 3 girls so the probability is $1/3$ ” or “Because there’s 2 female and 3 male.” One student who chose C, however, did so because, “There are 3 black puppies and 1 is a male.” This student, rather than being confused about the meaning of rational numbers, simply computed the wrong conditional probability (black instead of brown), and he/she did so correctly. This way of choosing C appeared to be an aberration rather than the pattern, so although C could be reached through a reasoning process not emerging from a misconception, the explanations of students in this sample indicated that the choice was overwhelmingly due to the misconception. No such aberrations appeared in explanations for Choice E. Therefore, both C and E were retained as a misconception response.

Students who chose D for Item 2 provided explanations that indicated two reasoning processes, both of which represented thinking based on misconceptions about rational numbers. The first explanation, used by the majority of students who chose D, relied on a comparison of brown dogs to dogs; these students made statements such as, “There are 2 male puppies and only 1 is brown” or “Because all together there is 1 male black and 1 brown, add them up which = 2 (so, $1/2$).” These responses indicated confusion about which quantities should have been represented by the part whole relationship. The second type of explanation relied instead on the uniformity heuristic as described by Falk (1992). These students chose D, “Because there are only 2 types of genders that you can pick.” Whether from confusion about part-whole

relationship quantities or the uniformity heuristic, the evidence from student explanations indicated that choice D did represent a misconception.

Item 3 response patterns. Item 3 also asked students about a conditional probability. In this scenario, Bill had a bag of 30 candies, 10 each of red, blue, and green. Using a random draw, Bill ate two pieces of blue candy. Students are then asked if the probability of getting a blue candy on the next draw is still $10/30$ or $1/3$. I hypothesized that students who chose A (yes) would ignore the conditional aspect of the probability altogether and that students who chose B (no) would recognize that the quantity making up the part-whole relationship had changed. Student explanations to both choices verified this hypothesis. Students who chose A sometimes relied on the uniformity heuristic, making statements such as, “Because there are 3 colors, and 1 could be picked blue.” Others who chose A simply ignored the conditional, justifying their response with statements such as, “You have 10 of each color candy to add up to $10/30$ ” or “Because there are ten and all together are 30.”

In contrast to explanations of choice A, students who chose B did not appear to do so by guessing or elimination. Indeed, these students recognized that the consumption of the two pieces of blue candy changed the quantities represented by the probability: “He ate $2/10$, so it’s now 8 blues instead of 10,” “Because he has already eaten 2, which lessens his chances,” “He already ate 2 of them, so its 28 left,” “Because his chances go down,” or “Because he took out 2 candies; it’s now $8/28$.” Based on these explanations, I concluded that students in this sample who chose correctly did indeed demonstrate a stronger conceptual understanding of the meaning of the rational number quantities present in the probability, so Choice A was retained as a misconception indicator for Item 3.

Item 4 response patterns. Item 4 presented students with a spinner divided in two halves,

with one of the halves divided in half again. The arrow on the spinner pointed to one of the quarter regions. Students were then asked how many times they should expect the arrow to land in that region after 300 spins. I hypothesized that a spatial reasoning misconception would result in students deciding that the probability of the region was $\frac{1}{3}$ instead of $\frac{1}{4}$. Based on this hypothesis, I expected Choice C to result from students taking $\frac{1}{3}$ of 300. I also expected students to choose D by taking $\frac{1}{3}$ of 360, the number of degrees of the circle. Explanations for choosing C verified that part of my hypothesis. These students made statements such as, “Because there’s 3 spaces and $300 \div 3$ is 100” or “I divided $300 \div 3$ & got 100.” Therefore, I retained Choice C as a misconception choice. Students who chose D, on the other hand, did not verify my hypothesis. Instead, these students indicated that they had recognized the probability of the region as being $\frac{1}{4}$ but made a computation error. No one who chose D gave explanations that indicated a misconception, so it was eliminated as a misconception choice. Explanations for choosing B did, however, indicate a misconception, possibly one of spatial reasoning or rational number meaning. These students recognized that the region was $\frac{1}{4}$ of the circle, so they fell back on the number of degrees in a circle. These students made statements such as, “The circle is in an angle of 90° ,” “Circle = 360, divide it by 4, you get 90,” or “Because the circle is split up into 3 parts; a circle’s measure is 360, if cut in half, each part will be 180, if one half of the split circle is cut in half again, that side is now 2 sets of 90° .” This error could be due to misunderstanding about the meaning of the quantities in a probability ratio, or it could be due to misunderstanding the quantity being predicted, focusing on a single circle instead of the same circle 300 times. Regardless of which misconception led to this choice, it seemed clear that choosing B for Item 4 represented at least one type of misconception, so it was added as a misconception choice for this Item. As with Item 16, the convergence of fundamental

misconceptions on multiple content areas increased the likelihood of a high degree of collinearity between content area misconceptions.

Item 17 response patterns. Item 17 asked students to visualize a cube whose faces are labeled R or S. The probability of landing on R was given as $1/3$. Students were then asked to determine how many faces of the cube should be labeled R. I hypothesized that Choices C and E would represent misconceptions about absolute/relative comparisons or the meaning of rational numbers. Students who chose these responses and explained their answers corroborated this expectation with statements such as “because $1/3 \rightarrow 1$,” “one because there is a one out of three chance,” or “the number on the bottom is how many.” Some students who chose C ignored the stated probability altogether, similar to students who ignored rational number quantities in Item 5 (as students did with variables in Küchemann, 1978). These students relied instead on the uniformity heuristic (Falk, 1992), stating that “each face R and S gets three sides” and “half of 6 is 3.” From these explanations, I concluded that C and E should be retained as representative of misconception-based reasoning.

Misconceptions on Rational Number Content Knowledge Items

Rational number items (i.e., Items 12, 13, 14, and 15; Appendix N) on the NAEP instrument examined student errors on rational numbers based on absolute/relative comparison, additive/multiplicative structure, and rational number meaning misconceptions.

Item 12 response patterns. Item 12 presented students with a situation in which the postage cost for a letter is based on a different rate for the first ounce. I hypothesized that Choice D would indicate an additive/multiplicative structure misconception (as in Moss et al., 2008; Warren, 2000). Explanations by students who chose D confirmed this hypothesis, making statements such as, “I multiplied $.33¢$ times 2.7” and “Because $33 + 33 = 66$; $0.7 \rightarrow 1 \rightarrow 66 + 22$ ”

= 88.” Student explanations of other choices revealed additional misconception responses for Item 12. Students who chose E followed the same reasoning as students who chose D, but they remembered to round the 2.7 to 3. So, these students wrote explanations such as “33 x 3 ounces = 99 cents.” Students who chose B used two types of reasoning to arrive at their answer. First, students added “33 + 22 + 11,” “33 + 22 + 0.7,” or “You have 2 whole ounces, $33\text{¢} + 22\text{¢} = 55\text{¢}$, next you have to figure out the .7. Take $22 \cdot 70\%$, which is 18. So you add 12 to 55, total would be 66¢ .” This final statement, apart from the readily apparent calculation errors, shows the same basic reasoning as the first two. This particular justification included several erasures over numbers, a characteristic that appeared to indicate that the student had changed numbers to arrive at the closest answer available. So, not only does the response demonstrate the same additive/multiplicative structure misconception, it may also indicate the persistence of the misconception even in the face of numbers not adding up correctly. The second type of justification for choice B relied on a multiplicative-only strategy, for example, “I multiplied 33 times 2.” Students who chose A used similar reasoning to that used by students who chose B. These students also dropped the 0.7 or, in one case, the second ounce. Most of these justifications were some variation of, “33 + 22 = 55.” The student who ignored the second ounce stated, “First ounce is 33 cents, next 0.7 of ounce is 22, rounded.” As a result of this analysis, I included choices A, B, and E as misconception responses in addition to the original hypothesized choice D. The presence of additive/multiplicative structure misconceptions in rational numbers as well as algebra emphasizes the likelihood of collinearity between content area misconceptions.

Item 13 response patterns. Item 13 presented students with a diagram and a scale of $\frac{3}{4}$ in = 10 ft. They were then asked to convert 48 feet to the scale drawing length. I hypothesized

that Choices D and E would represent reasoning indicative of rational number meaning misconceptions. Students chose D because, “ $3/4 \times 10 = 7.5$ in” or “ $3 \div 4 = 0.75$, 7.5 in.” These students demonstrated the first rational number misinterpretation identified by the analysis of Item 5 responses – performing the correct rational number operation to the wrong quantity (as in Fosnot & Dolk, 2002). Therefore, Choice D was retained as a misconception response.

Students who chose E, however, did not justify their choices with responses that implied that they had ignored the numerator as I had hypothesized. Instead, they made statements such as, “Because you add $3/4$ to 48 ft” or “Add them all up and divide by 10.” These responses clearly indicated an error in thinking, but I chose to discard E as a misconception choice because I could not find a clear connection between misconceptions identified by previous research and the reasoning process that led to this choice.

Students who chose C for Item 13 indicated a misconception that I did not anticipate in my hypotheses. These students indicated that they had ignored the rational number altogether, the fifth rational number mis-interpretation identified by the analysis of Item 5. Some students justified their choice with statements such as, “ $48 \div 10 = 4.8$ or 5.” Therefore, choice C was included as a misconception response for Item 13.

Item 14 response patterns. In Item 14, students were asked to arrange a set of three fractions in ascending order, distinguishing between absolute and relative comparisons. Students who responded with choice E recognized the relative relationship between fractions, but they misunderstood the nature of that relationship. For example, students understood that fractional numbers compare differently than other rational number forms, but they failed to recognize that the denominator quantity is the one that has this inverse relationship (as in Baturo, 1994; Behr et al., 1992; Lamon, 1999). Students who chose E simply stated that “smaller fractions are larger.”

Although this faulty reasoning clearly demonstrates a misunderstanding, these students did not demonstrate the absolute/relative comparison misconception, so E was not included as a misconception response. In contrast, students who chose B cited a comparison of the numerators only, ignoring the impact of the denominator on the overall quantity (i.e., absolute versus relative comparison misconception). Therefore, B was the only response recognized as a misconception response for Item 14.

Item 15 response patterns. Item 15 presented students with a rectangle divided into two columns of 10 cells for a total of 20 cells, six of which were shaded. Students were then asked to determine which rational number best represented the probability of the shaded region. I hypothesized that students who chose D would do so because they used a part-part relationship rather than a part-whole relationship. Explanations by students who chose D verified this hypothesis, making statements such as, “3 are black and 7 are white.” Therefore, Choice D was retained as a misconception choice for Item 15.

Implications of Item Response Patterns

This analysis of student responses to the NAEP mathematics knowledge test fundamentally altered the way misconception responses were coded. For several items, student responses validated the hypothesized misconception choices and rationales for each choice. For several other items, student responses suggested that the hypothesized misconception responses were either not due to misconceptions at all, not due to the hypothesized misconception, or not due to the hypothesized misconception for the correct reasons. As a result, the coding of these items was changed to match student response patterns to maximize content validity prior to the structural analysis of content area misconceptions and the analysis of contextual factors on the emergence of misconceptions on a particular item. Table 18 summarizes the changes from

hypothesized to observed misconception responses.

Table 18
Summary of Observed Misconception Responses

Item	Hypothesized Misconception Responses	Observed Misconception Responses
1	C, E	C, E
2	B, C, D, E	B, C, D, E
3	A	A
4*	C, D	B, C
5*	A, B, C	A, B, C, E
6*	D, E	B, C, D, E
7	A, B, C, D	A, B, C, D
8*	B	C, E
9	B	B
10*	D	C, D
11*	A	A, B, C
12*	D	A, B, D, E
13*	D, E	C, D, E
14	B	B
15	D	D
16*	A	A, C, D
17	C, E	C, E

*Indicates an item where coding was changed because of response analysis.

Three Level Bernoulli Model

The student level of the HLM model was then divided into two levels, item characteristics and student characteristics. By doing so, the outcome variable become a dichotomous variable representing a misconception error for each item for each student in each class. The new model was then examined using a generalized HLM model (HGLM) to measure the probability of misconception errors. The initial null model was examined to determine the amount of variance at each level: item level 1, student level 2, and class level 3. The contextual model was then used to evaluate the impact of each variable on the outcome.

Descriptive Statistics. The observed sample sizes (Table 19) resulted in a statistical power of approximately 0.80 to detect a population effect size $\delta = 0.40$ and approximately 0.75

for a population effect size $\delta = 0.30$ for approximately 20 students per class. In this sample, class sizes averaged approximately 18 students.

Table 19
Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Item Level One					
Misconception	9673	0.35	0.48	0.00	1.00
Discrimination	9673	0.80	0.20	0.44	1.22
Difficulty	9673	0.01	0.51	-1.21	0.96
Moderate	9673	0.35	0.48	0.00	1.00
Student Level 2					
Enjoyment	515	2.91	0.79	1.00	5.00
Motivation	515	2.90	0.93	1.00	5.00
Self Confidence	515	3.19	0.83	1.00	5.00
Value	515	3.51	0.75	1.00	5.00
Knowledge of Cognition	515	3.43	0.58	1.18	4.94
Regulation of Cognition	515	3.23	0.56	1.11	4.74
NAEP Pretest Percent Misconception	515	0.38	0.17	0.00	0.76
Class Level 3					
Mean Enjoyment	32	2.90	0.30	2.35	3.52
Mean Motivation	32	3.43	0.16	3.01	3.65
Mean Self Confidence	32	3.23	0.16	2.92	3.50
Mean Value	32	0.38	0.10	0.14	0.53
Mean Knowledge of Cognition	32	2.90	0.30	2.35	3.52
Mean Regulation of Cognition	32	2.88	0.34	2.15	3.63
Mean Pretest Percent Misconception	32	3.17	0.32	2.41	3.96
Treatment	32	0.50	0.51	0.00	1.00

Unconditional Null Model. The unconditional model (Equations 3, 4, and 5) revealed a significant amount of variance at both the student Level 2 and class Level 3 (Table 20). Additionally, the level 1 variance was also statistically significant (SE = 0.014, $t = 67.71$)

Item Level 1 (3)

$$Prob(Misconception_{ijk} = 1 | \pi_{jk}) = \phi_{ijk}$$

$$Log \left[\frac{\phi_{ijk}}{1 - \phi_{ijk}} \right] = \eta_{ijk}$$

$$\eta_{ijk} = \pi_{0jk} + e_{ijk}$$

Student Level 2 (4)

$$\pi_{0jk} = \beta_{0k} + r_{0jk}$$

Class Level 3 (5)

$$\beta_0 = \gamma_{00} + u_{0k}$$

Table 20

Unconditional Three Level Model Fixed and Random Coefficients

Fixed Effects	Logit Link: Unit-Specific Model		Logit Link: Population Average Model	
Mean Item Misconception, γ_{000}	-0.643**		-0.590**	
<i>Random Effects</i>	<i>Variance Component</i>	<i>df</i>	χ^2	<i>p Value</i>
Between Classes, u_{00}	0.184**	31	188.001	< 0.001
Between Students, R_0	0.299**	483	1117.002	< 0.001
Between Items, E	0.945**			

*Indicates |coeff/se| > 2.00; ** Indicates |coeff/se| > 3.00;

The outcome variable for Level 1 is in logit units, or the natural logarithm of the odds ratio, as shown in Equation 2. The coefficients, therefore, are also computed in logit units. Using the logit unit, the relationship between the outcome variable and independent variable coefficients have a linear relationship. Once the coefficient logit is converted to a probability, its relationship to the outcome variable and other logit coefficients is no longer linear. Therefore, to compute a predicted probability of misconception error for an item, the predicted logit value must be computed first. Conversion to a probability is the final step in predicting outcomes in the Bernoulli HLM model. The process of converting from a logit to a probability requires two steps. First, the logit is converted to an odds ratio using Equation 6.

$$\text{Odds} = e^{\text{Logit}} \quad (6)$$

Second, the inverse of the odds ratio is used to compute the probability using Equation 7.

$$\text{Probability} = \left(\frac{1}{1 + e^{-\text{Logit}}} \right) \quad (7)$$

The γ_{000} logit (unit specific model) of -0.643 (Table 93) corresponds to an odds ratio of 0.526, or a probability of 0.35 for the appearance of a misconception on an item, which corresponds to the mean for Item Misconception (see Table 19). The logit of the population average model (-0.590) indicates that the expected appearance of misconceptions in the population is slightly different from the observed sample misconception probability, an odds ratio of 0.554 and a probability of 0.357. This difference represents the expected effect of τ_{00} , in this case pulling the mean value of the unit specific model upward toward a probability of 0.50 (Raudenbush & Bryk, 2002).

The total variance in the model equals the sum of the variance from all three levels (Table 20), $0.184 + 0.299 + 0.945 = 1.428$. The proportion of variance at the item Level 1 is $0.184/1.428 = 0.129 = 12.9\%$. The proportion of variance at the student Level 2 is $0.299/1.428 = 0.209 = 20.9\%$. The proportion of variance at the class Level 3 is $0.945/1.428 = 0.662 = 66.2\%$. In summary, the variance at each level was statistically significant, and the class level 3 accounted for the majority of the variance in the probability of misconception errors. To begin accounting for variance, the item level 1 model was calibrated.

Item level 1 model. The discrimination and difficulty IRT coefficients for each NAEP item (see Table 4) were used as explanatory variables in the level 1 model. Additionally, the reported level of complexity assigned by NAEP reviewers (see Table 4) was added as a dichotomous predictor of misconception errors on a particular item (Low Complexity = 0;

Moderate Complexity = 1). Two models were examined before arriving at the final item model (Equations 8, 9, and 10; Table 21).

Item Level 1 (8)

$$Prob(Misconception_{ijk} = 1 | \pi_{jk}) = \phi_{ijk}$$

$$Log \left[\frac{\phi_{ijk}}{1 - \phi_{ijk}} \right] = \eta_{ijk}$$

$$\eta_{ijk} = \pi_{0jk} + \pi_{1jk}(\text{Discrimination}) + \pi_{2jk}(\text{Difficulty}) + \pi_{3jk}(\text{Complexity}) + e_{ijk}$$

Student Level 2 (9)

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk}$$

$$\pi_{2jk} = \beta_{20k} + r_{2jk}$$

$$\pi_{3jk} = \beta_{30k}$$

Class Level 3 (10)

$$\beta_{00k} = \gamma_{000} + u_{00k}$$

$$\beta_{10k} = \gamma_{100} + u_{10k}$$

$$\beta_{20k} = \gamma_{200} + u_{20k}$$

$$\beta_{30k} = \gamma_{300} + u_{30k}$$

The variance components for class level discrimination (U10), class level complexity (U30), and student level complexity (R3) were statistically non-significant, so they were fixed in the final item model. The addition of the discrimination, difficulty, and complexity variables reduced the item level variance from 0.945 to 0.881, a 6.8% reduction.

Table 21
Item Model of Mathematics Misconception Errors

	Unit Specific Model		Population Average Model	
	Unconditional Model ^b	Final Model	Unconditional Model ^b	Final Model
<i>Fixed Effects</i>				
Intercept, γ_{000}	-0.643**	0.233*	-0.590	0.130
Discrimination, γ_{100}	—	-1.206***	—	-0.957***
Difficulty, γ_{200}	—	0.682***	—	0.565***
Complexity, γ_{300}	—	0.080	—	0.094*
<i>Variance Components</i>				
Intercept U_{00} , $\tau(\beta)_{11}$	Unconditional Model ^b		Final Model	
	0.184***		0.195***	
Class Discrimination U_{10} , $\tau(\beta)_{22}$	—		—	
Class Difficulty, U_{20} , $\tau(\beta)_{33}$	—		0.056*	
Class Complexity, U_{30} , $\tau(\beta)_{44}$	—		—	
Std Intercept, R_0 , $\tau(\pi)_{11}$	0.299***		1.044***	
Std Discrimination, R_1 , $\tau(\pi)_{22}$	—		2.480***	
Std Difficulty, R_2 , $\tau(\pi)_{33}$	—		0.385***	
Std Complexity, R_3 , $\tau(\pi)_{44}$	—		—	
Item, E , σ^2	0.945 ^a		0.881 ^a	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ^a $|\text{coeff}/\text{SE}| > 3.00$

^b $df_{\text{Student}} = 483$, $df_{\text{Class}} = 31$

Using Equations 6 and 7, the logits for the final item model were converted into predicted probabilities of misconception errors for different item characteristics. The intercept logit value predicts that the probability of a misconception error for a non-discriminating item (i.e., item characteristic curve = horizontal line) of average difficulty (i.e., Difficulty = 0) and low complexity is 0.558 in the sample and 0.532 in the population (Table 100). If the difficulty of a non-discriminating item increases difficulty by one standard deviation (0.62), then the probability of a misconception error increases to 0.658 for the sample and 0.618 for the population (Table 22).

If an item has average discrimination (0.825, mean of discrimination values from Table 4), then the predicted probability of a misconception error reduces to 0.318 in the sample and 0.341 in the population. If the discrimination of an item has a value one standard deviation above the average discrimination ($0.8 + 0.2 = 1$), then the predicted probability of a misconception error reduces to 0.230 in the sample and 0.277 in the population (Table 22). If a non-discriminating item with an average difficulty level increases from low to moderate complexity,

the probability of a misconception error increases to 0.578 for the sample and 0.556 for the population (Table 22).

Table 22
Selected Predicted Values for Final Item Model

Fixed Effects	Unit Specific Model			Population Average Model		
	Logit	Odds	Probability	Logit	Odds	Probability
INTERCEPT, γ_{000}	0.234	1.263	0.558	0.130	1.139	0.532
Discrimination, γ_{100}	-1.206	0.299	0.230	-0.957	0.384	0.277
Difficulty, γ_{200}	0.682	1.978	0.664	-0.565	1.759	0.638
Complexity, γ_{300}	0.080	1.083	0.520	-0.094	1.099	0.523

Combined Effects						
Int + Mean Discrimination	-0.731	0.481	0.325	-0.636	0.530	0.346
Int + 1SD Above Mean Discrimination						
Int + 2SD Above Mean Discrimination	-1.696	0.183	0.155	-1.018	0.361	0.265
Int + 1SD Above Mean Difficulty	0.588	1.801	0.643	0.424	1.528	0.604
Int + 1SD Below Mean Difficulty	-0.107	0.898	0.473	-0.153	0.859	0.462
Int + Complexity	0.314	1.368	0.578	0.224	1.251	0.556

The probabilities and odds ratios for each logit value in Table 22 were computed using Equations 6 and 7. Combined effects were computed through a process of three steps. First, standard deviations of discrimination and difficulty were taken from Table 19. Second, the relevant number of standard deviations values were multiplied by the logit coefficient and added to the intercept logit. Third, the resulting logit sum was converted to an odds ratio and probability using Equations 6 and 7.

These predicted probability values reflect a statistically significant impact of item characteristics on the probability of a misconception error. The remaining variance of the item level was still statistically significant after the addition of all available variables, indicating that a future examination of other item characteristics may be beneficial to understanding item characteristic influences on misconception errors. The final item model was used as the starting point for calibration of the student model.

Student level 2 model. Student characteristics were added to level 2 (student level) of the final item model to examine the impact of student characteristics on the probability of a misconception error and on the impact of item characteristics on the probability of a misconception error. Only statistically significant effects were retained in the final model (Table 23) with the exception of the self confidence impact on the difficulty slope. Self confidence was retained because removing it from the model resulted in the loss of a significant coefficient for motivation, which was statistically significant in all intermediate models.

Table 23
Final Student Model

Fixed Effects	Logit Link:			
	Unit-Specific Model	Population Average Model		
Intercept, γ_{000}	0.237	0.165		
Discrimination Slope, γ_{100}	-1.221***	-1.027***		
Pretest Slope, γ_{110}	3.389***	2.961***		
Difficulty Slope, γ_{200}	0.689***	0.583***		
Motivation Slope, γ_{210}	-0.199*	-0.182*		
Self Confidence Slope, γ_{220}	0.163***	0.151		
Pretest Slope, γ_{230}	-0.799***	-0.525		
Complexity Slope, γ_{300}	-0.080	0.095*		
	<i>Variance</i>			
<i>Random Effects</i>	<i>Component</i>	<i>df</i>	χ^2	<i>p Value</i>
	Class Intercept, u_{00}	31	221.117	< 0.001
	Class Difficulty Slope	31	53.032	0.008
	Std Intercept, R_0	483	622.784	< 0.001
	Std Discrimination Slope, R_1	513	686.892	< 0.001
	Std Difficulty Slope, R_2	480	649.958	< 0.001
	Item Intercept, E		0.887***	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ^a|coeff/SE| > 3.00

As with the Item Model, predicted values are presented to clarify the meaning of coefficients computed as logits (Table 24). The process for computing these predicted probabilities was the same as for the Item Model (i.e., use Equations 6 and 7 to convert logits to probability). The fixed effect intercept, γ_{000} , represents a predicted probability of a misconception error on an item with no discrimination, average difficulty, and low complexity of 0.559 for the sample and 0.541 in the population. The fixed effect for discrimination, γ_{100} , means that the impact of a change of one logit unit in discrimination (for an item with average difficulty

and low complexity) corresponds to a probability of misconception error of 0.228 for the sample and 0.264.

Table 24
Selected Predicted Values for Final Student Model

Fixed Effects	Unit Specific Model			Population Average Model		
	Logit	Odds	Probability	Logit	Odds	Probability
Intercept, γ_{000}	0.237	1.267	0.559	0.165	1.179	0.541
Discrimination Slope, γ_{100}	-1.221	0.295	0.228	-1.027	0.358	0.264
Pretest Slope, γ_{110}	3.389	29.636	0.967	2.961	19.317	0.951
Difficulty Slope, γ_{200}	0.689	1.992	0.666	0.583	1.791	0.642
Motivation Slope, γ_{210}	-0.199	0.820	0.450	-0.182	0.834	0.455
Self Confidence Slope, γ_{220}	0.163	1.177	0.541	0.151	1.163	0.538
Pretest Slope, γ_{230}	-0.799	0.450	0.310	-0.525	0.592	0.372
Complexity Slope, γ_{300}	0.08	1.083	0.520	0.095	1.100	0.524
Combined Effects	$\hat{\eta}$	Odds	Probability	$\hat{\eta}$	Odds	Probability
Int + Mean Discrimination + Mean Difficulty + Low Complexity + Mean Pretest	-0.894	0.409	0.290	-0.862	0.422	0.297
Int + Mean Discrimination + Mean Difficulty + Low Complexity + 1SD Above Mean Pretest	-0.279	0.757	0.431	-0.254	0.776	0.437
Int + Mean Discrimination + 1SD Below Mean Difficulty + 1SD Above Mean Motivation + 1SD Above Mean Pretest	-0.174	0.840	0.457	-0.152	0.859	0.462
Int + Mean Discrimination + 1SD Above Mean Difficulty + 1SD Below Mean Motivation + 1SD Above Mean Pretest	0.677	1.968	0.663	0.613	1.846	0.649

The combined effects in Table 24 were computed by adding the relevant student logits to the item logits, then combining the item logits to produce the predicted logit for misconception errors ($\hat{\eta}$). These combined logits were then converted to probabilities using Equations 6 and 7. If a student with more pretest misconceptions (1 SD = 0.17) than the mean (0.38) completed an item with average discrimination (0.8), average difficulty (0), and low complexity, the probability of a misconception error increases to 0.431 in the sample and 0.437 in the population. If a student with more misconceptions (1 SD = 0.17) and more motivation (1 SD = 0.93) than the

mean (0.38, 2.9 respectively) completed an easy item (1 SD below difficulty mean = -0.5), the probability of a misconception error was 0.457 in the sample and 0.462 in the population. If a student with more misconceptions (1SD = 0.17) and less motivation (1 SD = 0.93) than the mean (0.38, 2.9 respectively) completed a difficult item (1 SD above difficulty mean = 0.502) with mean discrimination (0.8), the probability of a misconception error increases to 0.663 in the sample and 0.649 in the population.

Two student level slopes demonstrated statistically significant variance at Level 3 (class), the intercept U000 and motivation slope U200. The final student model was used as the initial model for calibrating class level variables.

Class level 3 model. Class characteristics were added as predictors to the two level 3 equations with statistically significant, the intercept (mean probability of misconception error) and the difficulty slope (impact of item difficulty on the probability of misconception error).

The addition of these class parameters produced the final model (Table 25), which reduced the class variance from 0.214 to 0.002, a 98.7% reduction. This reduction reflected a statistically significant reduction in level 3 model misfit ($\Delta\chi^2 = 188.482$, $\Delta df = 8$, $p < 0.001$).

Table 25
Final Class Model

Fixed Effects	Logit Link:			
	Unit-Specific Model	Population Average Model		
Intercept, γ_{000}	0.268*	0.200 ^b		
Class Enjoyment, γ_{001}	-0.885*	-0.867*		
Class Motivation, γ_{002}	-0.109	-0.079		
Class Self Confidence, γ_{003}	0.234	0.246		
Class Value, γ_{004}	0.649*	0.597*		
Class Knowledge of Cognition, γ_{005}	0.118	0.164		
Class Regulation of Cognition, γ_{006}	0.349	0.307		
Class Pretest, γ_{007}	4.676***	4.423***		
Treatment, γ_{008}	-0.051	-0.055		
Discrimination Slope, γ_{100}	-1.215***	-1.067***		
Pretest Slope, γ_{110}	3.379***	3.100***		
Difficulty Slope, γ_{200}	0.814***	-0.730***		
Class Enjoyment, γ_{201}	-1.384*	-1.319*		
Class Motivation, γ_{202}	0.338	0.290		
Class Self Confidence, γ_{203}	-0.961*	0.894***		
Class Value, γ_{204}	0.118	0.158**		
Class Knowledge of Cognition, γ_{205}	0.138**	0.090		
Class Regulation of Cognition, γ_{206}	0.350**	0.363		
Class Pretest, γ_{207}	-0.536	-0.257		
Treatment, γ_{208}	-0.231 ^b	-0.208 ^b		
Motivation Slope, γ_{210}	-0.199*	-0.187*		
Self Confidence Slope, γ_{220}	0.162	0.154		
Pretest Slope, γ_{230}	-0.792***	-0.583		
Complexity Slope, γ_{300}	0.080**	0.097*		
	<i>Variance</i>			
<i>Random Effects</i>	<i>Component</i>	<i>df</i>	χ^2	<i>p</i> Value
Class Intercept, u_{00}	0.002 ^b	23	32.635	0.088
Class Difficulty Slope, u_{20}	0.008*	23	35.125	0.050
Std Intercept, R_0	1.044***	483	621.070	< 0.001
Std Discrimination Slope, R_1	1.494***	513	683.774	< 0.001
Std Difficulty Slope, R_2	0.597***	480	650.180	< 0.001
Item Intercept, E	0.943 ^a			

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ^a|coeff/SE| > 3.00, ^b $p \leq 0.10$

The treatment condition was not a statistically significant predictor of the intercept (mean probability of misconception error), but it was a statistically significant predictor of the difficulty slope at the 90% confidence level for both the sample and population models. The coefficients from Table 25 were used to compute the predicted probability for a misconception error under various item, student, and class conditions (Table 26).

Table 26
Predicted Values for Final Class Model

Fixed Effects	Unit Specific Model			Population Average Model		
	Logit	Odds	Probability	Logit	Odds	Probability
Intercept, γ_{000}	0.268	1.307	0.567	0.200	1.221	0.550
Class Enjoyment, γ_{001}	-0.885	0.413	0.292	-0.867	0.420	0.296
Class Motivation, γ_{002}	-0.109	0.897	0.473	-0.079	0.924	0.480
Class Self Confidence, γ_{003}	0.234	1.264	0.558	0.246	1.279	0.561
Class Value, γ_{004}	0.649	1.914	0.657	0.597	1.817	0.645
Class Knowledge of Cognition, γ_{005}	0.118	1.125	0.529	0.164	1.178	0.541
Class Regulation of Cognition, γ_{006}	0.349	1.418	0.586	0.307	1.359	0.576
Class Pretest, γ_{007}	4.676	107.340	0.991	4.423	83.346	0.988
Treatment, γ_{008}	-0.051	0.950	0.487	-0.055	0.946	0.486
Discrimination Slope, γ_{100}	-1.215	0.297	0.229	-1.067	0.344	0.256
Pretest Slope, γ_{110}	3.379	29.341	0.967	3.100	22.198	0.957
Difficulty Slope, γ_{200}	0.814	2.257	0.693	0.730	2.075	0.675
Class Enjoyment, γ_{201}	-1.384	0.251	0.200	-1.319	0.267	0.211
Class Motivation, γ_{202}	0.338	1.402	0.584	0.290	1.336	0.572
Class Self Confidence, γ_{203}	0.961	2.614	0.723	0.894	2.445	0.710
Class Value, γ_{204}	0.118	1.125	0.529	0.158	1.171	0.539
Class Knowledge of Cognition, γ_{205}	0.138	1.148	0.534	0.090	1.094	0.522
Class Regulation of Cognition, γ_{206}	0.35	1.419	0.587	0.363	1.438	0.590
Class Pretest, γ_{207}	-0.536	0.585	0.369	-0.257	0.773	0.436
Treatment, γ_{208}	-0.231	0.794	0.443	-0.208	0.812	0.448
Motivation Slope, γ_{210}	-0.199	0.820	0.450	-0.187	0.829	0.453
Self Confidence Slope, γ_{220}	0.162	1.176	0.540	0.154	1.166	0.538
Pretest Slope, γ_{230}	-0.792	0.453	0.312	-0.583	0.558	0.358
Complexity Slope, γ_{300}	0.08	1.083	0.520	0.097	1.102	0.524

Table 26 (Continued)
Predicted Values for Final Class Model

Combined Effects	Unit Specific Model			Population Average Model		
	$\hat{\eta}$	Odds	Probability	$\hat{\eta}$	Odds	Probability
Int + Mean Discrimination + Mean Difficulty + Low Complexity + Mean Class Enjoyment, Self Confidence, Motivation + Class Mean Pretest + Std Mean Pretest + Std Mean Motivation + Control	-0.281	0.755	0.430	-0.274	0.760	0.432
Int + Mean Discrimination + Mean Difficulty + Low Complexity + Mean Self Confidence, Motivation + Class Mean Pretest + Std Mean Pretest + Std Mean Motivation + 1SD Above Mean Class Enjoyment + Control	-0.762	0.467	0.318	-0.740	0.477	0.323
Int + Mean Discrimination + Mean Difficulty + Low Complexity + 1SD Above Class Mean Self Confidence, Enjoyment, Motivation + Class Mean Pretest + Std Mean Pretest + Std Mean Motivation + Control	-0.201	0.818	0.450	-0.200	0.819	0.450
Int + Mean Discrimination + Mean Difficulty + Low Complexity + Class Mean Self Confidence, Enjoyment, Motivation, Pretest + Std Mean Pretest, Motivation + Treatment	-0.401	0.670	0.401	-0.382	0.682	0.406

The computation of probabilities from the fixed effects in Table 26 proceeded as in the item and student models, using Equations 6 and 7. The combined effects, however, required a consideration of the effects of class variables on student variables before combining student effects with item effects to produce the predicted value.

The first combined effect predicted the probability of a misconception error on an item of mean discrimination, mean difficulty, and low complexity for a student with mean pretest misconceptions and motivation in a control class with mean self confidence, enjoyment and motivation. Because all student and class level variables were centered (group and grand

centered respectively), the student and class mean values produced a zero effect on the predicted probability. The intercept logit for this combined effect was 0.268. The discrimination impact was -1.215 (the coefficient logit) \bullet 0.8 (the average discrimination) = -0.972 , the value added to the intercept logit. The difficulty impact was 0.814 , the coefficient logit \bullet 0.52 , 1 SD above average difficulty = 0.4233 , the value added to the intercept logit. Complexity was coded as a dichotomous variable in which low complexity was coded as 0. Therefore, the predicted logit, $\hat{\eta}$, for this situation was $0.268 + -0.972 + 0.4233 = -0.281$. The associated probability of misconception error (using Equations 6 and 7) was 0.430 in the sample. The same computational process was used for the population average model and subsequent combined effect examples.

The second and third combined effects from Table 26 represent the probability of a misconception error on a hard item (1 SD above mean Difficulty = 0.52) of mean discrimination (0.8) and low complexity (see Table 19 for difficulty and discrimination values). For combined effect 2, a student who had an average pretest misconception score in a control class with average pretest misconception, class enjoyment, motivation, value, and self confidence scores was predicted by the model to have a probability of a misconception error of 0.430 in the sample and 0.432 in the population. For the same student (Combined Effect 3), if the class enjoyment level increased by one standard deviation (0.3 units), the probability of a misconception error reduced to 0.318 in the sample and 0.323 in the population (Table 26).

The fourth combined effect examines the relationship of self confidence with misconception error probabilities. A student with pretest misconception score equal to the mean, on a hard item of average discrimination and low complexity, in a control class with average enjoyment, motivation, value, and pretest mean misconceptions but one standard deviation above

the mean for self confidence had a misconception error probability of 0.450 for both the sample and population (Table 26).

The treatment condition had an indirect effect on the probability of a misconception error by impacting the item difficulty slope. A student with a mean pretest misconception score and motivation in a treatment class with average enjoyment, self confidence, value, and pretest misconception was predicted to have a probability of misconception error of 0.401 in the sample and 0.406 in the population on an item of average discrimination and difficulty and low complexity (Table 26).

Discussion

Two analyses were conducted to investigate the nature of misconceptions in mathematics: (1) analysis of student response patterns on the mathematics knowledge test (NAEP items) and (2) examination of the impact of item, student, and class characteristics on misconception errors. Item characteristics were measured using Item Response Theory on the NAEP mathematics knowledge test. Student characteristics included attitudes toward mathematics (ATMI Enjoyment, Motivation, Self Confidence, and Value scales), metacognitive awareness (MAI Knowledge of Cognition and Regulation of Cognition scales), and pre- and post-test misconception and percent correct scores. Class characteristics consisted of aggregated scores for each student characteristic along with indicator variables for treatment condition and type of mathematics class. Through these two analyses, five key findings emerged.

1. Five underlying misconceptions affect all four content areas.
2. Mathematics misconception errors often appear as procedural errors.
3. A classroom environment that fosters enjoyment of mathematics and value of mathematics are associated with reduced misconception errors.

4. Higher mathematics self confidence and motivation to learn mathematics is associated with reduced misconception errors.
5. Probability instruction may not affect misconceptions directly, but it may help students develop skills needed to bypass misconceptions when solving difficult problems.

Analysis 1: Misconception Error Analysis

Two key findings. The first analysis of the present study presented patterns of student responses on the mathematics knowledge test composed of NAEP items. This document analysis validated hypotheses about how misconceptions would result in error choices for eight of the 17 items. Misconception error choices for the other nine items were adjusted to align with observed student responses (as shown in Table 18) before proceeding with the quantitative analyses. The observed patterns of misconception errors revealed an important aspect of mathematics misconceptions, Key Finding 1: On a wide array of mathematical problems, a very small number of fundamental misconceptions (five) appeared to account for a large proportion of the observed errors (70.49%). All five of these core misconceptions (i.e., Absolute/Relative Comparison, Additive/Multiplicative Structure, Spatial Reasoning, Variable Meaning, and Rational Number Meaning Misconceptions) appeared in multiple mathematics content areas.

Another conclusion emerged from the analysis of student response patterns, Key Finding 2: Misconception error explanations relied on procedural knowledge isolated from conceptual knowledge (as described in Figure 10). Previous studies (e.g., Agnoli & Krantz, 1989; Kahneman & Tversky, 1972; Falk, 1992) have also indicated that reliance on judgmental heuristics may be an important factor in the development of mathematics misconceptions.

Our task as mathematics educators is to distinguish between those circumstances in which judgmental heuristics can adversely affect stochastic

thinking and those in which the heuristics are useful and desirable. And we are obliged to point out the differences to our students. It is not that there is “something wrong” with the way our students think. It is just that they (and we) tend to carry useful heuristics beyond their relevant domain (Shaughnessy & Bergman, 1993, p. 184).

The analysis of the present study indicates that connecting procedural knowledge to conceptual knowledge may help teachers and students make these distinctions. Hiebert and Grouws (2007) described two observable features for a classroom that focuses on developing conceptual understanding: (1) Teaching focuses explicitly to connections between facts, procedures, and ideas, and (2) Students are allowed to struggle with important mathematical concepts. Development of these two features in a classroom may help teachers identify the reasoning behind errors that emerge from misconceptions.

Detecting mathematics misconceptions. NAEP released items were compiled “as is,” without any changes for the mathematics knowledge test (Appendix M). By doing so, the NAEP-established item content and concurrent criterion validity could be transferred to the present study (Daro et al., 2007). The compiled instrument also exhibited acceptable internal consistency and test-retest reliability. Despite these qualities, the instrument failed to adequately differentiate between misconceptions.

The ambiguity in student explanations for several items indicated that the validity of the items did not necessarily extend to measuring misconceptions. For example, the question of whether content area or type of underlying misconception category is a better way to organize mathematical misconceptions cannot be answered by the present study — some item responses indicated multiple types of misconceptions (e.g., Item 7, Choice D; Item 17, Choice C). Such a

question might be answerable using a multi-trait, multi-method structural equation model, but a model of this type, based on the present instrument would most likely require several cross-loadings that would make the model structurally unstable (i.e., no amount of iterations can yield a solution), such as was seen in Models A, B, and C of the structural analysis. Therefore, I recommend that such a study begin by altering the present instrument to focus directly on observed misconception responses. For example, in Item 17, the uniformity heuristic sometimes represented an absolute/relative comparison misconception, an additive/multiplicative structure misconception, a rational number meaning misconception, or a combination of these misconception types. To distinguish misconceptions more readily, it may be necessary to include explanations with possible answers. For example, instead of simply offering the choice “three,” a revised item might offer “one because the numerator is one (absolute/relative comparison misconception), “three because the denominator is three” (rational number meaning misconception), and “three because R and S have equal faces” (additive/multiplicative structure misconception via the uniformity heuristic). Without such differentiation, misconception content validity for closed-response items will be difficult to establish. A study to develop and validate such a misconception instrument may be a necessary first step to replicating and advancing the present investigation.

Analysis 2: Factors Influencing Misconception Errors

1. *Three key findings.* The final analysis of the present study examined the impact of item, student, and class characteristics on misconception errors. The results of this analysis led to Key Findings 3, 4, and 7: A classroom environment that fosters enjoyment of mathematics or value of mathematics helps reduce student misconception errors; Higher mathematics self confidence reduces misconception errors; and, Probability instruction

may not affect misconceptions directly, but it may help students develop skills needed to bypass misconceptions when solving difficult problems.

The two level model revealed significant predictors of misconceptions for both student and class characteristics. Student mathematics self confidence (ATMI self confidence scale) and pretest misconception error percentages (NAEP instrument) accounted for 29% of the student variance in posttest misconception error percentages (NAEP instrument). Class enjoyment of mathematics (ATMI enjoyment scale) and the class value of mathematics (ATMI value scale) also had a statistically significant effect on posttest misconception errors. The between-class variance in the unconditional model (see Table 89) was 0.0073 ($p < 0.001$). The contextual model that included the statistically significant class variables (see Table 92) reduced the between class variance to 0.0005 ($p = 0.016$). This reduction represented a 93.15% reduction. The three level model also accounted for 98% of the class level variance. Such large reductions in variance indicates that a large percentage of class effects on misconceptions may lie in the factors measured by the ATMI scales. If true, then educators can begin focusing on improvement of these factors within a class to reduce misconceptions.

Implications for teaching mathematics. Traditional mathematics instruction has relied primarily on teacher-centered epistemologies (Stigler & Hiebert, 1999). This investigation began with the assumption that student-centered instructional approaches have a more positive effect on student mathematics learning than traditional, teacher-centered strategies. The present study supported this assumption and extended it to addressing misconceptions. Higher mathematics self confidence, value, and enjoyment were associated with a decline in misconceptions; the development of a positive learning environment may therefore be a critical component to helping students traverse the learning barriers in Figure 35.

Maier and Tetrault (2001) described four epistemological components critical to developing such a positive learning environment: mastery, voice, authority, and positionality. First, mastery involves struggle and engagement with a body of knowledge. Instead of merely absorbing information, students grapple with difficulties of understanding. Rather than the final product as end goal, mastery refers to the continual process of working and re-working information into knowledge. Second, the fashioning of one's voice in mathematics means to bring one's personal experiences, questions, and perspectives to the mathematics being studied. Third, the concept of authority refers to the source of mathematical knowledge in a classroom. Maier and Tetrault (2001) and Shrewsbury (1993) described a climate of shared mathematical authority: Students and teachers share the knowledge and understanding of important mathematical ideas in such an environment. Authority refers to the relationship between students and teachers collectively with mathematical knowledge. Fourth, positionality refers to the relationships between an individual and mathematical knowledge along with the interactions of these within- and between-student relationships.

As teachers seek to help students turn the barriers of Figure 35 into opportunities to reinforce fundamental mathematics concepts, the development of a student-centered environment may be a foundational component for any strategy. Previous studies (e.g., Slavin & Karweit, 1982; Slavin & Lake, 2008; Slavin et al., 2009) have found that student-centered teaching approaches provide benefits to student achievement. If these environments are to offer the most benefit to avoiding and addressing misconceptions, then students must be given opportunities to struggle with important mathematical ideas and their connections (Hiebert & Grouws, 2007; Kieran, 1989, 1992, 2007).

An ontological perspective from the present study also offers insight for helping students

overcome the learning barriers in Figure 35. The examination of student misconception error explanations revealed a consistent pattern: Misconception errors occurred when students relied on procedures isolated from meaning and mathematical structure. This pattern suggests that mathematics is best understood as an organized structure of meanings and connections rather than procedures.

Teachers should strive to organize the mathematics so that fundamental ideas form an integrated whole. Big ideas encountered in a variety of contexts should be established carefully, with important elements such as terminology, definitions, notation, concepts, and skills emerging in the process (NCTM, 2000, p. 14)

In combination with the epistemological implications described above, the present study found that student-centered, concept-focused mathematics classrooms may be the most effective learning environment for turning fundamental mathematics barriers into opportunities to learn.

Conclusions

Future investigations of mathematics misconceptions may best begin by developing a more refined instruments for identifying and categorizing misconceptions. The present study explored the use of probability instruction as an intervention to target fundamental mathematics misconceptions. The treatment condition had a statistically significant impact on the effect of item difficulty on misconception error probabilities, and several other important statistically significant factors were also identified. Only five foundational concepts appeared to be fundamental to learning mathematics. Attending to these five foundational concepts may allow mathematics teaching in any single area to fundamentally impact the learning of other mathematics areas. Ignoring this small set of foundational concepts may allow the development

of a formidable obstruction at a level that can inhibit and perhaps derail the mathematics future of students. Such an astounding notion may indicate that finding ways to identify and address these foundational concepts and their connections to a particular mathematics area should be one of the primary, critical tasks for mathematics educators.

DRAFT

REFERENCES

- Adamson, J. W. (Ed.). (1922). *The educational writings of John Locke*. Cambridge: University Press.
- Agnoli, F. (1987). Teaching implications of misconceptions in probability and statistics. In J. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. I, pp. 1-5). Ithaca, NY: Cornell University.
- Agnoli, F., & Krantz, D. H. (1989). Suppressing natural heuristics by formal instruction: The case of the conjunction fallacy. *Cognitive Psychology*, *21*, 515-550.
- Ahadi, S. A., & Rothbart, M. K. (1994). Temperament, development, and the Big Five. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 189-207). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ainley, J., Bills, L., & Wilson, K. (2004, July). *Constructing meanings and utilities within algebraic tasks*. Paper presented at the 28th conference of the International Group for the Psychology of Mathematics Education, Bergen, Norway.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.
- Akaike, H. (1983). Information measures and model selection. *Bulletin of the International Statistical Institute: Proceedings of the 44th Session*, *1*, 277-290.
- Alomar, B. O. (2007). Personal and family factors as predictors of pupils' mathematics achievement. *Psychological Reports*, *101*, 259-269.
- Alpine Media Corporation. (2003). *The Temperament Sorter II: A statistical study*. Orem, UT: Author.
- Amenkhienan, C. A., & Kogan, L. R. (2004). Engineering students' perceptions of academic activities and support services: Factors that influence their academic performance. *College Student Journal*, *38*, 523-540.
- Amir, R., Frankl, D. R., & Tamir, P. (1987). Justifications of answers to multiple choice items as a means for identifying misconceptions. In J. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. I, pp. 15-25). Ithaca, NY: Cornell University.
- Anderson, S. E. (1990). Worldmath curriculum: Fighting Eurocentrism in mathematics. *Journal of Negro Education*, *59*, 348-359.
- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, *48*, 12-19.
- Ashman, S. M. (2001). Consumer choice models with customer loyalty programs in retail food stores. *Dissertation Abstracts International – A*, *61*(11), 4484. (UMI No. AAT 9994493)
- Baker, F. B., & Kim, S.-H. (2004). *Item response theory: Parameter estimation techniques*. New York: Marcel Dekker.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*, 389-409.
- Bandalos, D. L. (1993). Factors influencing cross-validation of confirmatory factor analysis models. *Multivariate Behavioral Research*, *28*, 351-374.
- Bandalos, D. L. (2009, April). *A comparison of fit index behavior, parameter estimate, and standard error bias for some mean- and variance-adjusted estimators*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Bar-Hillel, M., & Falk, R. (1982). Some teasers concerning conditional probabilities. *Cognition*, *11*, 109-122.

- Barkatsas, A., Kasimatis, K., & Gialamas, V. (2009). Learning secondary mathematics with technology: Exploring the complex interrelationship between students' attitudes, engagement, gender and achievement. *Computers & Education*, *52*, 562-570.
- Baroudi, Z. (2006). Easing students' transition to algebra. *Australian Mathematics Teacher*, *62*, 28-33.
- Barry, R. R. (1988). The application of sabermetrics to the teaching and coaching of collegiate baseball. (Doctoral dissertation, Middle Tennessee State University, 1988). *Dissertation Abstracts International-A*, *50*(02), 384.
- Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 843-908). Reston, VA: National Council of Teachers of Mathematics.
- Baturo, A. R. (1994, August). *A cognitive analysis of misconceptions in year 12 students' understanding of elementary probabilistic notions*. Paper presented at the 3rd International Seminar of Misconceptions and Educational Strategies in Science and Mathematics, Ithaca, NY.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions (1683-1775)*, *53*, 370-418.
- Behr, M. (1980). How children view the equal sign. *Mathematics Teaching*, *92*, 13-15.
- Behr, M. J., Harel, G., Post, T. R., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 296-333). Reston, VA: National Council of Teachers of Mathematics.
- Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, *62*, 750-762.
- Bell, A. (1987). Misconceptions, conflict and discussion in the teaching of graphical interpretations. In J. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. I, pp. 46-58). Ithaca, NY: Cornell University.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.
- Berg, K. F. (1993, April). *Structured cooperative learning and achievement in a high school mathematics class*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Blanco, L., & Garrote, M. (2007). Difficulties in learning inequalities in students of the first year of pre-university education in Spain. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*, 221-229.
- Bloodbook.com. (2007). *Blood type facts*. Retrieved on May 14, 2009, from <http://www.bloodbook.com/type-facts.html>
- Boland, P. J., & Nicholson, J. (1996). The statistics and probability curriculum at the secondary school level in the USA, Ireland and the UK. *The Statistician*, *45*, 437-446.
- Bong, M. (2008). Effects of parent-child relationships and classroom goal structures on motivation, help-seeking avoidance, and cheating. *Journal of Experimental Education*, *76*, 191-217.
- Booth, L. R. (1986). Difficulties in algebra. *Australian Mathematics Teacher*, *42*, 2-4.
- Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, *40*, 119-156.
- Bowles, S., & Levin, H. M. (1968). The determinants of scholastic achievement — An appraisal of some recent evidence. *The Journal of Human Resources*, *3*, 3-24.
- Brandt, A. E., & Pietras, C. J. (2008). Gambling on a simulated slot machine under conditions of repeated play. *The Psychological Record*, *58*,

- 405-426.
- Brick, J. M., Jones, M. E., Kalton, G., & Valliant, R. (2005). Variance estimation with hot deck imputation: A simulation study of three methods. *Survey Methodology*, *31*, 151-159.
- Briggs, J., Demana, F., & Osborne, A. R. (1986). Moving into algebra: Developing the concepts of variable and function. *Australian Mathematics Teacher*, *42*, 5-8.
- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*, 445-455.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park, CA: Sage.
- Brownell, W. A. (1935). Psychological considerations in the learning and teaching of arithmetic. In W. D. Reeve (Ed.), *The teaching of arithmetic: Tenth yearbook of the National Council of Teachers of Mathematics* (pp. 1-31). New York: Teachers College, Columbia University.
- Brownell, W. A. (1956). Meaning and skill: Maintaining the balance. *Arithmetic Teacher*, *34*, 18-25.
- Bryk, A. S., & Raudenbush, S. W. (1988). Heterogeneity of variance in experimental studies: A challenge to conventional interpretations. *Psychological Bulletin*, *104*, 396-404.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Byrne, B. M. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Capaldi, D. M., & Rothbart, M. K. (1992). Development and validation of an early adolescent temperament measure. *Journal of Early Adolescence*, *12*, 153-173.
- Capraro, M. M., Kulm, G., & Capraro, R. M. (2005). Middle grades: Misconceptions in statistical thinking. *School Science and Mathematics*, *105*, 165-174.
- Carraher, D. W., & Schliemann, A. D. (2007). Early algebra. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 669-706). Reston, VA: National Council of Teachers of Mathematics.
- Carrell, S. E., Page, M. E., & West, J. E. (2009, May). *Sex and science: How professor gender perpetuates the gender gap* (Working Paper No. 14959). Cambridge, MA: National Bureau of Economic Research. (ERIC Document Reproduction Service No. ED505322)
- Carter, J. A., Cuevas, G. J., Day, R., Malloy, C., Cummins, J., & Zike, D. (2010). *Geometry*. Columbus, OH: Glencoe/McGraw-Hill.
- Carter, S. (2008). Disequilibrium and questioning in the primary classroom: Establishing routines that help students learn. *Teaching Children Mathematics*, *15*, 134-137.
- Castro, C. S. (1998). Teaching probability for conceptual change. *Educational Studies in Mathematics*, *35*, 233.
- Catanzaro, D. (1997). Course enrichment and the job characteristics model. *Teaching of Psychology*, *24*, 85-87.
- Chambers, D. L. (1994). The right algebra for all. *Educational Leadership*, *53*, 85-86.
- Chang, Y. (2002, March). *Understanding and learning of function: Junior high school students in Taiwan*. Paper presented at the International Conference on Mathematics – “Understanding Proving and Proving to Understand,” Taipei, Taiwan.

- Chazan, D., & Yerushalmy, M. (2003). On appreciating the cognitive complexity of school algebra: Research on algebra learning and directions of curricular change. In J. Kilpatrick, W. G. Martin, & D. Schifter, *A research companion to "Principles and standards for school mathematics"* (pp. 123-135). Reston, VA: National Council of Teachers of Mathematics.
- Clement, L. L. (2001). What do students really know about functions? *Mathematics Teacher*, *94*, 745-748.
- Clement, J. (1982). Algebra word problem solutions: Analysis of a common misconception. *Journal for Research in Mathematics Education*, *13*, 16-30.
- Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 420-464). Reston, VA: National Council of Teachers of Mathematics.
- Clotfelter, C. T., & Cook, P. J. (1991). Lotteries in the real world. *Journal of Risk and Uncertainty*, *4*, 227-232.
- Cobb, P., Gresalfi, M., & Hodge, L. L. (2009). An interpretive scheme for analyzing the identities that students develop in mathematics classrooms. *Journal for Research in Mathematics Education*, *40*, 40-68.
- Coes, L. (1993). Building fractal models with manipulatives. *Mathematics Teacher*, *86*, 646-651.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Mahweh, NJ: Lawrence Erlbaum Associates.
- Coleman, J. S. (1967a). 'Equality of educational opportunity,' reconsidered. Washington, DC: U.S. Department of Health, Education & Welfare Office of Education. (ERIC Document Reproduction Service No. ED015893)
- Coleman, J. S. (1967b). *The concept of 'Equality of educational opportunity.'* Washington, DC: U.S. Department of Health, Education & Welfare Office of Education. (ERIC Document Reproduction Service No. ED015157)
- Coleman, J. S. (1968). *The evaluation of 'Equality of educational opportunity.'* Washington, DC: U.S. Department of Health, Education, & Welfare Office of Education. (ERIC Document Reproduction Service No. ED026721)
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington DC: U.S. Department of Health, Education & Welfare Office of Education. (ERIC Document Reproduction Service No. ED012275)
- CollegeBoard Advanced Placement. (2008). *Statistics course description*. New York: CollegeBoard.com Inc.
- Collis, K. F. (1975). *A study of concrete and formal operations in school mathematics: A Piagetian viewpoint*. Hawthorn, Victoria: Australian Council for Educational Research.
- Confrey, J. (1987). 'Misconceptions' across subject matters: Science, mathematics and programming. In J. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. I, pp. 81-106). Ithaca, NY: Cornell University.
- Confrey, J. (2006, November 11). Fuzzy policy, not 'Fuzzy math,' Is the problem. *Education Week*. Retrieved on January 19, 2009, from <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=23096977&site=ehost-live>
- Confrey, J., & Lipton, A. (1985). Misconceptions research and the clinical interview. In S. K. Damarin & M. Shelton (Eds.), *Proceedings of the Seventh Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 38-43), Columbus, OH.
- Cooney, T. J., Davis, E. J., & Henderson, K. B. (1975). *Dynamics of teaching secondary school mathematics*. Boston: Houghton Mifflin.
- Coppen, M. J. (1976). The development and evaluation of an individualized mastery programme in mathematics for low-ability secondary school

- students. *New Zealand Journal of Educational Studies*, 11, 132-142.
- Crawford, M., Souders, J., & Pedrotti, L. (2009). *Cord algebra 1: Learning in context*. Waco, TX: Cord Communications.
- Creswell, J. W. (2007). *Qualitative inquiry and research design* (2nd ed.). Thousand Oaks, CA: Sage.
- Crowley, M. L. (1987). The Van Hiele model of the development of geometric thought. In M. M. Lindquist & A. P. Shulte (Eds.), *Learning and teaching geometry, K-12: 1987 yearbook* (pp. 1-16). Reston, VA: National Council of Teachers of Mathematics.
- Daggett, W. R. & Kruse, B. (1999). *Taming the educational dinosaur*. Rexford, NY: Leadership Press.
- Daro, P., Stancavage, F., Ortega, M., DeStefano, L., & Linn, R. (2007). *Validity study of the NAEP mathematics assessment: Grades 4 and 8*. Washington, DC: American Institutes for Research. Retrieved March 28, 2010, from http://www.air.org/publications/documents/NAEP_Math_Validity_Study.pdf
- Davis, B. (2007). Learners within contexts that learn. In T. Lamberg & L. R. Wiest (Eds.), *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* Stateline, NV: University of Nevada, Reno.
- Davis, R. B. (1984). *Learning mathematics: The cognitive science approach to mathematics education*. Norwood, NJ: Ablex.
- Davis, R. B. (1992). Reflections on where mathematics education now stands and on where it may be going. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 724-734). Reston, VA: National Council of Teachers of Mathematics.
- Deal or No Deal. (2009). *Deal or no deal game show: Play the TV game show online*. New York: NBC.com. Retrieved October 1, 2009, from http://www.nbc.com/Deal_or_No_Deal/game/flash.shtml
- De Bock, D., Van Dooren, W., Verschaffel, L., & Janssens, D. (2002). Improper use of linear reasoning: An in-depth study of the nature and the irresistibility of secondary school students' errors. *Educational Studies in Mathematics*, 50, 311-334.
- De Bock, D., Verschaffel, L., & Janssens, D. (1998). The predominance of the linear model in secondary school students' solutions of word problems involving length and area of similar plane figures. *Educational Studies in Mathematics*, 35, 65-83.
- Dermitzaki, I., Leondari, A., & Goudas, M. (2009). Relations between young students' strategic behaviours, domain-specific self-concept, and performance in a problem-solving situation. *Learning and Instruction*, 19, 144-157.
- Dietiker, L., Kysh, J., Hoey, B., & Salley, T. (2006). *Algebra connections* (Rev. ed.). Sacramento, CA: College Preparatory Mathematics.
- Dietiker, L., Kysh, J., Hoey, B., & Sallee, T. (2007). *Geometry connections* (Version 3.1). Sacramento, CA: CPM Educational Program.
- Douglas, J. B. (1970). On the teaching of probability and statistics at the pre-college level in Australia. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 303-318). Stockholm: Almqvist & Wiksell Förlag AB.
- Edwards, T. G. (2000). Some 'Big ideas' of algebra in the middle grades. *Mathematics Teaching in the Middle School*, 6, 26-31.
- Ellis, L. K., & Rothbart, M. K. (2001, April). *Revision of the Early Adolescent Temperament Questionnaire*. Poster session presented at the Biennial Meeting of the Society for Research in Child Development, Minneapolis, Minnesota. Retrieved July 11, 2009, from <http://www.bowdoin.edu/~sputnam/rothbart-temperament-questionnaires/pdf/lesa-ellis-srcd-poster-reprint.pdf>
- Enfedaque, J. (1990). De los números a las letras [From numbers to letters]. *Suma*, 5, 23-34.
- Enger, N. (2006). Relationship between mobility and student performance and behavior. *Journal of Educational Research*, 99, 167-178.
- Engel, A. (1970). Teaching probability in intermediate grades. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 87-150). Stockholm: Almqvist & Wiksell Förlag AB.

- Esposito Lamy, C. (2003). Academic achievement in the context of poverty: Examining the effects of family risk on children's math and reading achievement over the elementary school years. *Dissertation Abstracts International-A*, 64(04), 1172. (UMI No. AAT 3088472)
- Evans, J., & Tsatsaroni, A. (2000). Mathematics and its publics: Texts, contexts and users. *Social Epistemology*, 14, 55-68.
- Falk, R. (1981). The perception of randomness. In C. Comiti, *Psychology of mathematics education. Proceedings of the fifth conference of the international group* (pp. 222-229). Seoul, South Korea. (ERIC Document Reproduction Service No. ED 225809)
- Falk, R. (1992). A closer look at the probabilities of the notorious three prisoners. *Cognition*, 43, 197-223.
- Falk, R., & Konold, C. (1994). Random means hard to digest. *Focus on Learning Problems in Mathematics*, 16, 2-12.
- Falk, R., & Lann, A. (2008). The allure of equality: Uniformity in probabilistic and statistical judgment. *Cognitive Psychology*, 57, 293-334.
- Fajemidagba, O. (1983). The relationships between Piagetian cognitive developmental stages of concrete and formal operations and achievement on mathematical ratio and proportion problems. (Doctoral dissertation, State University of New York at Albany). *Dissertation Abstracts International-A*, 43(11), 3497.
- Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Education and Psychological Measurement*, 53, 357-382.
- Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and affective factors. *American Educational Research Journal*, 14, 51-71.
- Fennema, E., & Sherman, J. (1978). Sex related differences in mathematics achievement and related factors: A further study. *Journal for Research in Mathematics Education*, 9, 189-203.
- Festinger, L. (1957). *Theorie der kognitiven Dissonanz [A theory of cognitive dissonance]*. Evanston, IL: Row, Peterson.
- Fey, J. T. (1979). Mathematics teaching today: Perspectives from three national surveys. *Mathematics Teacher*, 72, 490-504.
- Finn, C. E. (2005, January 26). Grading school standards: A tale of two studies. *National Review Online*, 57. Retrieved on June 29, 2009, from <http://www.nationalreview.com/comment/finn200501260740.asp>
- Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics*, 15, 1-24.
- Fisher, K. (1988). The student-and-professor problem revisited. *Journal for Research in Mathematics Education*, 19, 260-262.
- Ford, C. B. (1977). A study of the relationships among pupil perception of teacher social power base, teacher pupil control ideology and teacher concern level in teachers of secondary mathematics. (Doctoral dissertation, University of Houston, 1976). *Dissertation Abstracts International-A*, 37(08), 5051.
- Fosnot, C. T., & Dolk, M. (2002). *Young mathematicians at work: Constructing fractions, decimals, and percents*. Portsmouth, NH: Heinemann.
- Fraser, B. J. (1998). Classroom environment instruments: Development, validity, and applications. *Learning Environments Research*, 1, 7-33.
- Freeman, M. (1997). *Math and science on a personal level*. Lexington, KY: University of Kentucky, Appalachian Rural Systemic Initiative.
- Freudenthal, H. (1970). The aims of teaching probability. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 151-168). Stockholm: Almqvist & Wiksell Förlag AB.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Boston: Kluwer Academic.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Boston: Kluwer Academic.
- Fuson, K. C. (1998). Pedagogical, mathematical, and real-world conceptual-support nets: A model for building children's multidigit domain

- knowledge. *Mathematical Cognition*, 4, 147-186.
- Fuson, K. C., Kalchman, M. S., & Bransford, J. D. (2005). Mathematical understanding: An introduction. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 217-256). Washington DC: National Academies Press.
- Fuys, D. J., & Liebov, A. K. (1997). Concept learning in geometry. *Teaching Children Mathematics*, 3, 248-251.
- Garcia-Coll, C. T., Kagan, J., & Reznick, J. S. (1984). Behavioral inhibition in young children. *Child Development*, 55, 1005-1019.
- Gardner, H. (1987). The theory of multiple intelligences. *Annals of Dyslexia*, 37, 19-35.
- Gardner, H. (1989). Beyond a modular view of mind. In W. Damon (Ed.), *Child development today and tomorrow* (pp. 222-239). San Francisco, CA US: Jossey-Bass.
- Gardner, H. (2009). The five minds for the future. *School Administrator*, 66, 16-21.
- Gardner, H., & Hatch, T. (1989). Multiple intelligences go to school: Educational implications of the theory of multiple intelligences. *Educational Researcher*, 18, 4-9.
- Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist*, 24, 143-158.
- Garson, G. D. (2009). *Data imputation for missing values* [Website]. Raleigh, NC: North Carolina State University. Retrieved January 4, 2010, from <http://faculty.chass.ncsu.edu/garson/PA765/missing.htm>
- Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology*. (3rd ed.). Needham, MA: Allyn & Bacon.
- Gleick, J. (1987). *Chaos*. New York: Viking.
- Goffin, R. D. (1993). A comparison of two new indices for the assessment of fit of structural equation models. *Multivariate Behavioral Research*, 28, 205-214.
- Goldman, J., & Gardner, H. (1989). Multiple paths to educational effectiveness. In D.K.Lipsky, A. Gartner, D. K. Lipsky, & A. Gartner (Eds.), *Beyond separate education: Quality education for all* (pp. 121-139). Baltimore, MD England: Paul H. Brookes Publishing.
- Graham, A. T., & Thomas, M. O. J. (2000). Building a versatile understanding of algebraic variables with a graphic calculator. *Educational Studies in Mathematics*, 41, 265-282.
- Gray, S. S., Loud, B. J., & Sokolowski, C. P. (2005, October). *Undergraduates' errors in using and interpreting variables: A comparative study*. Paper presented at the 27th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Roanoke, VA.
- Green, D. R. (1982). *Probability concepts in 11-16 year old pupils*. Loughborough, England: Centre for Advancement of Mathematical Education in Technology, University of Technology. Retrieved on June 20, 2009, from <http://www.data-archive.ac.uk/doc/1946/mrdoc/pdf/a1946uab.pdf>
- Green, D. R. (1983a). From thumbtacks to inference. *School Science and Mathematics*, 83, 541-551.
- Green, D. R. (1983b). School pupils' probability concepts. *Teaching Statistics*, 5, 34-42.
- Gregg, J. (1995). Discipline, control, and the school mathematics tradition. *Teaching and Teacher Education*, 11, 579-593.
- Gresham, G., Sloan, T., & Vinson, B. (1997). *Reducing mathematics anxiety in fourth grade "at-risk" students*. Orlando, FL: University of Central Florida. (ERIC Document Reproduction Service No. ED417931)

- Ham, S., & Walker, E. (1999). *Getting to the right algebra: The Equity 2000 initiative in Milwaukee Public Schools*. Retrieved January 30, 2009, from <http://www.mdrc.org/publications/98/workpaper.html>
- Hatch, T. C., & Gardner, H. (1986). From testing intelligence to assessing competencies: A pluralistic view of intellect. *Roeper Review*, 8, 147-150.
- Haver, W. E. (1978). Developing skills in college algebra – A mastery approach. *The Two-Year College Mathematics Journal*, 9, 282-287.
- Hegedus, S. J., & Kaput, J. J. (2004, July). *An introduction to the profound potential of connected algebra activities: Issues of representation, engagement and pedagogy*. Paper presented at the annual meeting of the International Group for the Psychology of Mathematics Education, Bergen, Norway.
- Henderson, R. W., & Landesman, E. M. (1995). Effects of thematically integrated mathematics instruction on students of mexican descent. *Journal of Educational Research*, 88, 290-300.
- Hennequin, P. L. (1970). Teaching of probability and statistics in the French lycée. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 319-336). Stockholm: Almqvist & Wiksell Förlag AB.
- Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to "Principles and standards for school mathematics"* (pp. 5-23). Reston, VA: National Council of Teachers of Mathematics.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-100). Reston, VA: National Council of Teachers of Mathematics.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H. et al. (1997). *Making sense: Teaching and learning mathematics with understanding*. Portsmouth, NH: Heinemann.
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Reston, VA: National Council of Teachers of Mathematics.
- Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M. et al. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis*, 27, 111-132.
- Hoffman, E. S., & Caniglia, J. (2009). In their own words: Good mathematics teachers in the era of NCLB. *Mathematics Teacher*, 102, 468-473.
- Holliday, B., Cuevas, G. J., Moore-Harris, B., Carter, J. A., Marks, D., Casey, R. M. et al. (2005). *Algebra 2*. New York: Glencoe McGraw-Hill.
- Howe, R. (2005). *Comments on NAEP algebra problems*. Washington, D. C.: Brookings Institution. Retrieved on May 15, 2009, from <http://www.brookings.edu/brown/~media/Files/Centers/bcep/AlgebraicReasoningConferenceHowe.pdf>
- Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.
- Hughes, J., & Kwok, O. (2007). Influence of student-teacher and parent teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology*, 99, 39-51.
- Huinker, D., Doyle, L. H., & Pearson, G. E. (1995). *Landscape of mathematics and science education in Milwaukee. A study of the Milwaukee Public Schools*. Milwaukee, WI: University of Wisconsin-Milwaukee, Center for Mathematics and Science Education Research. (ERIC Document Reproduction Service No. ED379160)
- Hume, B. (1970). The introduction of probability and statistics on the pre-college level in Western Australia. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 337-340). Stockholm: Almqvist & Wiksell Förlag AB.

- Ismail, N. A. (2009). Understanding the gap in mathematics achievement of Malaysian students. *Journal of Educational Research*, 102, 389-394.
- Izsak, A. (2004). Students' coordination of knowledge when learning to model physical situations. *Cognition and Instruction*, 22, 81-128.
- Jendraszek, P. A. (2008). Misconceptions of probability among future teachers of mathematics. *Dissertation Abstracts International-A*, 69(01), 2008. (UMI No. AAT 3299270)
- Jones, F. H. (1987). *Positive classroom discipline*. Columbus, OH: McGraw-Hill.
- Jöreskog, K. G., & Sörbom, D. (1984). *LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Mooresville, IN: Scientific Software.
- Judson, T. W. (1999). Japan: A different model of mathematics education. *Contemporary Issues in Mathematics Education*, 36, 75-81.
- K-12 Mathematics Curriculum Center. (2005). *Curriculum summaries*. Author.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Kahneman, D., & Tversky, A. (1973a). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Kahneman, D., & Tversky, A. (1973b). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Kahneman, D., & Tversky, A. (1982). Subjective probability: A judgment of representativeness. In D. Kahneman, P. Slovic, & A. Tversky, *Judgement under uncertainty: Heuristics and Biases* (pp. 32-48). Cambridge: University Press.
- Kahneman, D., & Tversky, A. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293-315.
- Kalchman, M. S., & Koedinger, K. R. (2005). Teaching and learning functions. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 351-396). Washington D.C.: National Academies Press.
- Kant, E. (1781/1901). *Kritik der reinen Vernunft [Critique of pure reason]*. (J. M. D. Meiklejohn, Trans.). New York: P. F. Collier and Son.
- Kaput, J. J., & Clement, J. (1979). Letter to the editor. *Journal of Children's Mathematics Behavior*, 2, 208.
- Keirsey, D. (1998). *Please understand me II: Temperament, character, intelligence* (Revised ed.). Del Mar, CA: Prometheus Nemesis.
- Kentucky Department of Education. (2006). *Combined curriculum document*. Frankfort, KY: Kentucky Department of Education.
- Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Kieran, C. (1980, August). *The interpretation of the equals sign symbol for an equivalent relation vs. an operator symbol*. Paper presented at the Fourth International Conference for the Psychology of Mathematics Education, Berkely, CA.
- Kieran, C. (1989). The early learning of algebra: A structural perspective. In S. Wagner & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra* (pp. 33-56). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390-419). Reston, VA: National Council of Teachers of Mathematics.
- Kieran, C. (2007). Learning and teaching of algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 707-762). Reston, VA: National Council of Teachers of Mathematics.
- Kieran, C. (2008). *What do students struggle with when first introduced to algebra symbols?* Reston, VA: National Council of Teachers of Mathematics. Retrieved on June 22, 2009, from http://www.nctm.org/uploadedFiles/Research_News_and_Advocacy/Research/Clips_and_Briefs/Brief%20-

%20What%20Can%20We%20Learn.pdf

- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington D.C.: National Academy Press.
- Konold, C. (1988). *Understanding Students' Beliefs about Probability*. Massachusetts University: Amherst Scientific Reasoning Research Institute.
- Kozol, J. (1992). *Savage inequalities*. New York: Harper Perennial.
- Kozol, J. (2005). *The shame of the nation*. New York: Crown Publishers.
- Küchemann, D. (1978). Children's understanding of numerical variables. *Mathematics in School*, 9, 23-26.
- Küchemann, D. (1981). Cognitive demand of secondary school mathematics items. *Educational Studies in Mathematics*, 12, 301-316.
- Lai-Yin, L., & Rob, R. (2005). Chinese and English probabilistic thinking and risk taking in gambling. *Journal of Cross-Cultural Psychology*, 36, 621-627.
- Lamon, S. J. (1999). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers*. Philadelphia: Lawrence Erlbaum Associates.
- Lamon, S. J. (2007). Rational numbers and proportional reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629-667). Reston, VA: National Council of Teachers of Mathematics.
- LaRocque, M. (2008). Assessing perceptions of the environment in elementary classrooms: The link with achievement. *Educational Psychology in Practice*, 24, 289-305.
- Larson, R., Boswell, L., Kanold, T. D., & Stiff, L. (2010). *Algebra 2*. Austin, TX: Holt McDougal.
- Larson, R., & Farber, B. (2009). *Elementary statistics: Picturing the world*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Le, C. (2005, January 6). Latest report praises school standards – Accountability, climate rate high. *The News Journal*. Retrieved June 29, 2009, from http://infoweb.newsbank.com/iw-search/we/InfoWeb?p_product=AWNB&p_theme=aggregated5&p_action=doc&p_docid=1077FD9EDD2F927B&d_place=NWJB&f_subsection=sLOCAL&f_issue=2005-01-06&f_publisher=
- Lehrer, R., Strom, D., & Confrey, J. (2002). Grounding metaphors and inscriptional resonance: Children's emerging understanding of mathematical similarity. *Cognition & Instruction*, 20, 359-398.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60, 1-64.
- Leitzel, J. R. (1989). Critical considerations for the future of algebra instruction or a reaction to: "Algebra: what should we teach and how should we teach it?". In S. Wagner & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra* (pp. 25-32). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Leitzel, J. R., Ellis, W., Enneking, M., Heid, M. K., & Jones, P. L. (1991). *A call for change: Recommendations for the mathematical preparation of teachers of mathematics. An MAA report*. Washington, D.C.: Mathematical Association of America. (ERIC Document Reproduction No. ED412109)
- Leiva, M. A., Coes, L., & Brown, R. G. (2006). *Algebra 1: Explorations and applications*. Boston: Houghton Mifflin.
- Lin, X., Schwartz, D. L., & Hatano, G. (2005). Toward teachers' adaptive metacognition. *Educational Psychologist*, 40, 245-255.

- Liu, Y., & Thompson, P. (2007). Teachers' understandings of probability. *Cognition & Instruction*, 25, 113-160.
- Lopez, E. M., Gallimore, R., Garnier, H., & Reese, L. (2007). Preschool antecedents of mathematics achievement of Latinos: The influence of family resources, early literacy experiences, and preschool attendance. *Hispanic Journal of Behavioral Sciences*, 29, 456-471.
- Ma, X., Ma, L., & Bradley, K. D. (2008). Using multilevel modeling to investigate school effects. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59-110). Charlotte, NC: Information Age Publishing.
- MacGregor, M., & Stacey, K. (1993). Cognitive models underlying students' formation of simple linear equations. *Journal for Research in Mathematics Education*, 24, 217-232.
- MacGregor, M., & Stacey, K. (1997). Students' understanding of algebraic notation: 11-15. *Educational Studies in Mathematics*, 33, 1-19.
- Maher, F. A., & Tetrault, M. K. T. (2001). *The feminist classroom* (2nd ed.). Oxford: Rowman & Littlefield.
- Manouchehri, A., & Goodman, T. (1998). Mathematics curriculum reform and teachers: Understanding the connections. *Journal of Educational Research*, 92, 27-41.
- Mathematics, Science, & Technology Education. (2005). *The Cliff-Hanger*. Champagne, IL: University of Illinois. Retrieved July 1, 2009, from <http://mste.illinois.edu/activity/cliff/>
- Mathews, S. M. (1997). The effect of using two variables when there are two unknowns in solving algebraic word problems. *Mathematics Education Research Journal*, 9, 122-135.
- McAvoy, M. R. (2001). Essays in country banking: Case studies of the First National Bank of Bloomington (Illinois). *Dissertation Abstracts International-A*, 62(06), 2198. (UMI No. AAT 3017161)
- McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55, 503-513.
- McGatha, M. B., Bush, W. S., & Rakes, C. R. (2009). The effects of professional development in formative assessment on mathematics teaching performance and student achievement. *Journal of Multidisciplinary Evaluation*, 6, 32-43.
- McKinley, J. E. (1960). Relationship between selected factors and achievement in a unit on probability and statistics for twelfth grade students. (Doctoral dissertation, University of Pittsburgh, 1960). *Dissertation Abstracts*, 21(03), 561.
- McLeod, D. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). Reston, VA: National Council of Teachers of Mathematics.
- McMahon, A. J. (1979). The relationship between the pupil control ideology of mathematics teachers and student attitude toward mathematics. *Dissertation Abstracts International-A*, 40(05), 2524. (UMI No. AAT 7923654)
- McMahon, S. D., Wernsman, J., & Rose, D. S. (2009). The relation of classroom environment and school belonging to academic self-efficacy among urban fourth- and fifth-grade students. *Elementary School Journal*, 109, 267-281.
- Mendick, H. (2008). Subtracting difference: Troubling transitions from GCSE to AS-level mathematics. *British Educational Research Journal*, 34, 711-732.
- Ministry of Education (2007). *Secondary mathematics syllabuses*. Singapore: Author. Retrieved on March 28, 2009 from <http://www.moe.gov.sg/education/syllabuses/sciences/files/maths-secondary.pdf>
- Mitchell, C. E. (1990). Seven areas neglected in school math. *Education Digest*, 56, 64-67.
- Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives*. New York: Pantheon.

- Monaghan, F. (2000). What difference does it make? Children's views of the differences between some quadrilaterals. *Educational Studies in Mathematics*, 42, 179-196.
- Morgan, E., & Huebner, A. (2008). *Adolescent growth and development* (Publication No. 350-850). Blacksburg: Virginia Polytechnic Institute and State University. Retrieved on March 26, 2009, from <http://www.ext.vt.edu/pubs/family/350-850/350-850.pdf>
- Moschkovich, J. N. (1998). Resources for refining mathematical conceptions: Case studies in learning about linear functions. *Journal of the Learning Sciences*, 7, 209-237.
- Moss, J. (2005). Pipes, tubes, and beakers: New approaches to teaching the rational-number system. In M. S. Donovan & J. D. Bransford (Eds.), *How students learn: History, mathematics, and science in the classroom* (pp. 309-350). Washington D.C.: National Academies Press.
- Moss, J., Beatty, R., Barkin, S., & Shillolo, G. (2008). "What is your theory? What is your rule?" Fourth graders build an understanding of functions through patterns and generalizing problems. In C. E. Greenes & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics: Seventieth yearbook*. (pp. 155-168). Reston, VA: National Council of Teachers of Mathematics.
- Mueser, P. R., & Granberg, D. (1999). *The Monty Hall Dilemma revisited: Understanding the interaction of problem definition and decision making*. Unpublished manuscript, University of Missouri – Columbia. Retrieved on June 18, 2009, from <http://129.3.20.41/eps/exp/papers/9906/9906001.html>
- Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105, 430-445.
- Muraki, E., & Bock, D. (2002). *PARSCALE 4.1* [Computer Program]. Chicago: Scientific Software International.
- Murayama, K., & Elliot, A. J. (2009). The joint influence of personal achievement goals and classroom goal structures on achievement-relevant outcomes. *Journal of Educational Psychology*, 101, 432-447.
- Myers, I. B. (1962). *The Myers-Briggs type indicator*. Palo Alto, CA: Consulting Psychologists Press.
- 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 Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through Grade 8 mathematics: A quest for coherence*. Reston, VA: Author. Retrieved on September 5, 2008, from <http://www.nctm.org/standards/content.aspx?id=270>
- National Sciences Foundation, Mathematical Sciences Education Board. (1995). *Mathematical preparation of the technical workforce: Report of a workshop*. Washington, D. C.: National Academy Press. Retrieved on May 31, 2009, from http://www.nap.edu/catalog.php?record_id=9066#toc
- Nelson, R. D., Joseph, G. G., & Williams, J. (1993). *Multicultural mathematics: Teaching mathematics from a global perspective*. Oxford: Oxford University Press.
- Nemirovsky, R., & Ferrara, F. (2009). Mathematical imagination and embodied cognition. *Educational Studies in Mathematics*, 70, 159-174.
- Nichols, J. D., & Miller, R. B. (1994). Cooperative learning and student motivation. *Contemporary Educational Psychology*, 19, 167-178.
- No Child Left Behind Act of 2002, 20 U.S.C. § 6301 et seq. (West, 2002). Retrieved on February 1, 2009, from <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>
- Oberdorf, C. D., & Taylor-Cox, J. (1999). Shape up! *Teaching Children Mathematics*, 5, 340-345.
- Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research*, 70, 323-367.

- Palarea Medina, M. M. (1999). La adquisición del lenguaje algebraico: Reflexiones de una investigación [The acquisition of algebraic language: Reflections of an investigation]. *Números*, 40, 3-28.
- Papps, K. L. (2008). Productivity spillovers within families and firms. (Doctoral dissertation, Cornell University, 2008). *Dissertation Abstracts International*, 69(02), 2008.
- Parham, J. W. (1993). An analysis of the effects of tutoring on seventh-grade students engaged in the mastery of pre-algebra concepts. (Doctoral dissertation, University of Maryland College Park, 1993). *Dissertation Abstracts International-A*, 54(11), 4021.
- Pascal, B., & Fermat, P. (1654/2006). Fermat and Pascal on probability (V. Sanford, Trans.). In P. Tannery and C. Henry, *Oeuvres de Fermat* (pp. 288-314). New York: Simon & Schuster.
- Patton, M. W. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Philipp, R. A. (1992). A study of algebraic variables: Beyond the student-professor problem. *Journal of Mathematical Behavior*, 11, 161-176.
- Pólya, G. (1957). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press.
- Posner, M. I., & Rothbart, M. K. (2007). *Educating the human brain*. Washington D.C.: American Psychological Association.
- Presmeg, N. C. (1997, October). *A semiotic framework for linking cultural practice and classroom mathematics*. Paper presented at the meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Bloomington, IL.
- Protection of Human Subjects, 45 C.F.R. pt. 46. (2002). Retrieved June 1, 2009, from <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm>
- Quinn, R. J. (2004). Investigating probabilistic intuitions. *Teaching Statistics: An International Journal for Teachers*, 26, 86-88.
- Råde, L. (1970). Teaching of probability and statistics on the pre-college level in Sweden. In L. Råde (Ed.), *The teaching of probability and statistics* (pp. 341-346). Stockholm: Almqvist & Wiksell Förlag AB.
- Rakes, C. R., Valentine, J. C., & McGatha, M. B. (2010). *Methods of instructional improvement in algebra: A systematic review and meta-analysis*. Manuscript under review.
- Raudenbush, S. W. (2009). *Optimal Design* [Software]. Lincolnwood, IL: Scientific Software International.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2009). *Hierarchical linear modeling 6.08 for Windows* [Computer Software]. Lincolnwood, IL: Scientific Software International.
- Resnick, L. B. (1983). Mathematics and science learning: A new conception. *Science*, 220, 477-478.
- Resnick, M. D. (1987). *Choices: An introduction to decision theory*. Minneapolis: University of Minnesota Press.
- Reys, B. (2006). *The intended mathematics curriculum as represented in state-level curriculum standards: Consensus or confusion?* Charlotte, NC: IAP.
- Reys, B., & Lappan, G. (2007). Consensus or confusion? *Phi Delta Kappan*, 88, 676-680.
- Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, 91, 175-189.
- Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93, 346-362.

- Rodgers, C. E. (1995). An investigation of two instructional methods for teaching selected pre-algebra concepts to minority at-risk seventh-grade mathematics students. *Dissertation Abstracts International-A*, 56(08), 3042. (UMI No. AAT 9539938)
- Ronau, R. N., Rakes, C. R., Wagener, L., & Dougherty, B. (2009, February). *A comprehensive framework for teacher knowledge: Reaching the goals of mathematics teacher preparation*. Paper presented at the annual meeting of the Association of Mathematics Teacher Educators, Orlando, FL.
- Ronau, R. N., & Rakes, C. R. (2010). *A comprehensive framework for teacher knowledge: A lens for examining research and teaching*. Manuscript in preparation.
- Ronau, R. N., Wagener, L., & Rakes, C. R. (2009, April). A comprehensive framework for teacher knowledge: A lens for examining research. In R. N. Ronau (Chair), *Knowledge for Teaching Mathematics, a Structured Inquiry*. Symposium conducted at the annual meeting of the American Educational Research Association, San Diego, CA.
- Rosnick, P., & Clement, J. (1980). Learning without understanding: The effect of tutoring strategies on algebra misconceptions. *Journal of Mathematical Behavior*, 3, 3-27.
- Rothbart, M. K. (1994a). Broad dimensions of temperament and personality. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 337-341). New York: Oxford University Press.
- Rothbart, M. K. (1994b). Emotional development: Changes in reactivity and self-regulation. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 369-375). New York: Oxford University Press.
- Rothbart, M. K. (2006). *The Early Adolescent Temperament Questionnaire – Revised (EATQ-R)*. Retrieved July 11, 2009, from Bowdoin College, Mary Rothbart Temperament Lab Web site: <http://www.bowdoin.edu/~sputnam/rothbart-temperament-questionnaires/instrument-descriptions/early-adolescent-temperament.html>
- Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology*, 78, 122-135.
- Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly*, 40, 21-39.
- Rothbart, M. K., & Jones, L. B. (1998). Temperament, self-regulation, and education. *School Psychology Review*, 27, 479-492.
- Rouche, N. (2003). *Un fil conducteur pour l'enseignement des mathématiques [A thread for teaching mathematics]*. Nevelles, Belgium: Centre de Recherche sur l'Enseignement des Mathématiques [Center for Research on Teaching Mathematics]. Retrieved on July 9, 2009, from http://www.enseignement.be/download.php?do_id=2279&do_check=
- Rousseau, D. M. (1976). *The relationship of task characteristics of attitudes, absenteeism, stress, and performance among college students*. Washington, DC: U. S. Department of Health, Education, & Welfare National Institute of Education. (ERIC Document Reproduction Service No. ED140181)
- Rousseau, J.-J. (1979). *Emile: or, On education*. (A. Bloom, Trans.). New York: Basic Books. (Original work published 1762)
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: John Wiley & Sons.
- Rudasill, K. M. (2009). *Temperamental attention and activity, classroom emotional support, and academic achievement in third grade*. Manuscript submitted for publication.
- Rueda, M. R., Rothbart, M. K., Saccomanno, L., & Posner, M. I. (2007). Modifying brain networks underlying self-regulation. In D. Romer & E. F. Walker, *Adolescent psychopathology and the developing brain* (pp. 401-419). New York: Oxford University Press.

- Saul, M. (2008). Algebra: The mathematics and the pedagogy. In C. E. Greenes & R. Rubenstein (Eds.), *Algebra and algebraic thinking in school mathematics* (pp. 63). Reston, VA: National Council of Teachers of Mathematics.
- Schild, M. (2002). Statistical literacy inventory: Reading and interpreting tables and graphs involving rates and percentages. Minneapolis, MN: Augsburg College, W. M. Keck Statistical Literacy Project. Retrieved on June 19, 2009, from <http://web.augsburg.edu/~schild/MiloPapers/StatLitKnowledge2r.pdf>
- Schild, M. (2005). *Five percentage table survey* (Version 2b). Retrieved July 9, 2009, from Augsburg College, W. M. Keck Statistical Literacy Project Web site: <http://www.statlit.org/pdf/2005Schild5PctgTableSurvey.pdf>
- Schild, M. (2006). Statistical literacy survey results: Reading graphs and tables of rates and percentages. Minneapolis, MN: Augsburg College, W. M. Keck Statistical Literacy Project. Retrieved on June 19, 2009, from <http://www.augsburg.edu/statlit/pdf/2006SchildIASSIST.pdf>
- Schmidt, W., Houang, R., & Cogan, L. (2002). A coherent curriculum: The case of mathematics. *American Educator*, 26, 1-18.
- Schmidt, W. H., McKnight, C. C., Raizen, S. A., Jakwerth, P. M., Valverde, G. A., Wolfe, R. G., Britton, E. D., Bianchi, L. J., & Houang, R. T. (1997). A splintered vision: An investigation of U.S. science and mathematics education. Executive summary. *Wisconsin Teacher of Mathematics*, 48, 4-9.
- Schoenfeld, A. H. (1982, March). *Beyond the purely cognitive: Metacognition and social cognition as driving forces in intellectual performance*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Schoenfeld, A. H. (1985, April). *Students' beliefs about mathematics and their effects on mathematical performance: A questionnaire analysis*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Schoenfeld, A. H. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, 20, 338-355.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-371). Reston, VA: National Council of Teachers of Mathematics.
- Schoenfeld, A. H., Smith, J. L., & Arcavi, A. (1993). Learning: The microgenetic analysis of one student's evolving understanding of a complex subject matter domain. In R. Glaser (Ed.), *Advances in Instructional Psychology* (Vol. 4, pp. 55-175). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.
- Schreiber, J. B. (2000, October). *Scoring above the international average: A logistic regression model of the TIMSS advanced mathematics exam*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Chicago, IL.
- Schumacker, R. E., Young, J. I., & Bembry, K. L. (1995). Math attitudes and achievement of Algebra I students: a comparative study of computer-assisted and traditional lecture methods of instruction. *Computers in the Schools*, 11, 27-33.
- Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 623-660). Reston, VA: National Council of Teachers of Mathematics.
- Selvin, S. (1975a). Letters to the editor: A problem in probability. *American Statistician*, 29, 67-71.
- Selvin, S. (1975b). Letters to the editor: On the Monty Hall problem. *American Statistician*, 29, 131-134.

- Serra, M. (2003). *Discovering geometry: An investigative approach* (3rd ed.). Emeryville, CA: Key Curriculum Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental design*. Boston; New York: Houghton Mifflin Company.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M. (2003). Research on students' understandings of probability. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 216-226). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M. (2006). Research on students' understanding of some big concepts in statistics. In A. Flores, C. Pateracki, F. Bullard, J. M. Landwehr, & A. Hess (Eds.), *Thinking and reasoning with data and chance: 68th NCTM yearbook* (pp. 77-98). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M. (2007). Research on Statistics Learning and Reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M., & Bergman, B. (1993). Thinking about uncertainty: Probability and statistics. In P. S. Wilson (Ed.), *Research ideas for the classroom* (pp. 177-197). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M., & Burger, W. F. (1985). Spadework prior to deduction in geometry. *Mathematics Teacher*, 78, 419-428.
- Shaughnessy, J. M., & Dick, T. (1991). Monty's dilemma: Should you stick or switch? *The Mathematics Teacher*, 84, 252-256.
- Shermer, M. (2008). A random walk through middle land. *Scientific American*, 299, 40-41.
- Shevlin, M., & Miles, J. N. V. (1998). Effects of sample size, model specification, and factor loadings on the GFI in confirmatory factor analysis. *Personality & Individual Differences*, 25, 85-90.
- Shrewsbury, C. M. (1993). What is feminist pedagogy? *Women's Studies Quarterly*, 3-4, 8-15.
- Shulte, A. P. (1967). The effects of a unit in probability and statistics on students and teachers of ninth-grade general mathematics. (Doctoral dissertation, University of Michigan, 1967). *Dissertation Abstracts - A*, 28 (12), 4962.
- Sierpiska, A., & Kilpatrick, J. (1998). *Mathematics education as a research domain*. Kluwer Academic Publishers: Dordrecht.
- Sinclair, R., & Ghory, W. (1979, April). *Curriculum as environments for learning: A practical meaning and model*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Silver, H. F., Brunsting, J. R., & Walsh, T. (2008). *Math tools grades 3-12: 64 ways to differentiate instruction and increase student engagement*. Thousand Oaks, CA: Corwin Press.
- Silver, H., Strong, R., & Perini, M. (1997). Integrating learning styles and multiple intelligences. *Educational Leadership*, 55, 22-27.
- Skemp, R. R. (1976/2006). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20-26. Reprinted in *Mathematics Teaching in the Middle School*, 12, 88-95.
- Slavin, R. E., & Karweit, N. L. (1982). *Student teams and mastery learning: A factorial experiment in urban math nine classes*. Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools. (ERIC Document Reproduction Service No. ED215904)
- Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research*, 78, 427-515.

- Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79, 839-911.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3, 115-163.
- Smith, R. J. (2003). Equal arcs, triangles, and probability. *Mathematics Teacher*, 96, 618-621.
- Socas Robayna, M. M. (1997). Dificultades, obstáculos y errores en el aprendizaje de las matemáticas en la educación secundaria. [Difficulties, obstacles, and errors in learning mathematics in secondary education.]. In L. R. Romero (Ed.), *La educación matemática en la enseñanza secundaria. [Mathematics education at the secondary level.]* (pp. 125-154). Barcelona: Horsori y Instituto de Ciencias de la Educación.
- Soper, D. (2004). *Statistics calculators*. Retrieved February 20, 2009, from <http://www.danielsoper.com/statcalc/calc11.aspx>.
- Sowder, J. (1991). *Understanding as a Basis for Teaching: Mathematics and Science for Prospective Middle School Teachers. Final Report*. San Diego State University: Center for Research in Mathematics, Science, & Education. (ERIC Document Reproduction No. ED339590)
- Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction*, 18, 141-179.
- Spybrook, J. (2008). Power, sample size, and design. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 273-311). Charlotte, NC: Information Age Publishing.
- Stacey, K. (1989). Finding and using patterns in linear generalising problems. *Educational Studies in Mathematics*, 20, 147-164.
- Steen, L. A. (2007). How mathematics counts. *Educational Leadership*, 65, 8-14.
- Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling: A Multidisciplinary Journal*, 5, 411-419.
- Steiger, J. H., & Lind, J. C. (1980, June). *Statistically-based tests for the number of common factors*. Paper presented at the Psychometric Society Annual Meeting, Iowa City, IA.
- Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika*, 50, 253-264.
- Stein, Z., Connell, M., & Gardner, H. (2008). Exercising quality control in interdisciplinary education: Toward an epistemologically responsible approach. *Journal of Philosophy of Education*, 42, 401-414.
- Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319-369). Reston, VA: National Council of Teachers of Mathematics.
- Stemler, S. E., Elliott, J. G., Grigorenko, E. L., & Sternberg, R. J. (2006). There's more to teaching than instruction: Seven strategies for dealing with the practical side of teaching. *Educational Studies*, 32, 101-118.
- Stevens, J. (2001). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stibel, J. M., Dror, I. E., Ben-Zeev, T. (2009). The collapsing choice theory: Dissociating choice and judgment in decision making. *Theory and Decision*, 66, 149-179.
- Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instructions: An overview of the TIMSS video study. *Phi Delta Kappan*, 79, 14-21.

- Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge: Harvard University Press.
- Stiff, L. V., Johnson, J. L., & Johnson, M. R. (1993). Cognitive issues in mathematics education. In P. S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics* (pp. 3-20). Reston, VA: National Council of Teachers of Mathematics.
- Stipek, D. (2006). Relationships matter. *Educational Leadership*, 64, 46-49.
- Stone, J. R., Alfeld, C., & Pearson, D. (2008). Rigor and relevance: Enhancing high school students' math skills through career and technical education. *American Educational Research Journal*, 45, 767-795.
- Stringfield, S. C., & Teddlie, C. (2004). The multiple roles of N.L. Gage in the creation of educational effectiveness fields. *Teaching and Teacher Education*, 20, 397-404.
- Strong, R. W., Perini, M., Silver, H., & Thomas, E. (2004). Creating a differentiated mathematics classroom. *Educational Leadership*, 61, 73-78.
- Strong, R. W., Silver, H. F., & Perini, M. J. (2001). Making students as important as standards. *Educational Leadership*, 59, 56-61.
- Strutchens, M. (1995). *Multicultural mathematics: A more inclusive mathematics*. ERIC Digest. Washington D. C.: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED380295)
- Swaminathan, H. (2003). Bargaining and intra-household resource allocation: An analysis of the impact of credit and land in Malawi. *Dissertation Abstracts International-A*, 64(07), 2592. (UMI No. AAT 3098250)
- Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, 82, 306-314.
- Swenson, K. A. (1998). Middle school mathematics teachers' subject matter knowledge and pedagogical content knowledge of probability: Its relationship to probability instruction. *Dissertation Abstracts International-A*, 59(02), 440. (UMI No. AAT 9824767)
- Swindal, D. N. (2000). Learning geometry and a new language. *Teaching Children Mathematics*, 7, 246-250.
- Tankersley, M., Landrum, T. J., & Cook, B. G. (2004). How research informs practice in the field of emotional and behavioral disorders. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 98-116). New York: Guilford Press.
- Tapia, M., & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, 8, 16-21.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Tenenbaum, G. (1986). The effect of quality of instruction on higher and lower mental processes and on the prediction of summative achievement. *Journal of Educational Research*, 80, 105-114.
- Thomas, A., Chess, S., Birch, H. G., Hertzog, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.
- Thorndike, E. L. (1913). *Educational psychology* (Vols. 1-2). New York: Columbia University Press.
- Thorpe, J. A. (1989). Algebra: What should we teach and how should we teach it? In S. Wagner & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra* (pp. 11-24). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Torigoe, E., & Gladding, G. (2006). Same to us, different to them: Numeric computation versus symbolic representation. *AIP Conference Proceedings*, 883, 153-156.
- Tungate, M. (2008). Educators throughout state get 'thoughtful' about instruction. *Kentucky Teacher*, August 2008, 8-9.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105-110.
- Urbina, S. (2004). *Essentials of psychological testing*. Hoboken, NJ: John Wiley & Sons.

- U. S. Census Bureau. (2009). *Geography & environment: Environmental industry; Threatened and endangered species*. Washington, D. C.: The 2009 Statistical Abstract, U. S. Fish & Wildlife Service. Retrieved on May 14, 2009, from http://www.census.gov/compendia/statab/cats/geography_environment/environmental_industry_threatened_and_endangered_species.html
- U. S. Department of Education. (1996). *National Assessment of Educational Progress (NAEP) questions*. Retrieved on July 10, 2009, from <http://nces.ed.gov/nationsreportcard/itmrls/>.
- U. S. Department of Education. (2005). *National Assessment of Educational Progress (NAEP) questions*. Retrieved on November 3, 2008, from <http://nces.ed.gov/nationsreportcard/itmrls/>
- U. S. Department of Education. (2007). *National Assessment of Educational Progress (NAEP) questions*. Retrieved on November 3, 2008, from <http://nces.ed.gov/nationsreportcard/itmrls/>
- U. S. Department of Education. (2008a). *National Assessment of Educational Progress (NAEP) questions*. Retrieved on November 3, 2008, from http://nces.ed.gov/nationsreportcard/tdw/item_development/develop.asp
- U. S. Department of Education. (2008b). *Overall descriptive statistics for item blocks, as defined after scaling, grade 4 mathematics combined national and state main assessment: 2003*. Retrieved on May 26, 2009, from http://nces.ed.gov/nationsreportcard/tdw/analysis/2002_2003/initial_classical_descrip2003mathg4.asp
- U. S. Department of Education. (2008c). *Overall descriptive statistics for item blocks as defined after scaling, grade 8 mathematics combined national and state main assessment: 2003*. Retrieved on May 26, 2009, from http://nces.ed.gov/nationsreportcard/tdw/analysis/2002_2003/initial_classical_descrip2003mathg8.asp
- U. S. Department of Education. (2008d). *Overall descriptive statistics for item blocks as defined after scaling, grade 12 mathematics national main assessment (accommodations not permitted [R2]): 2000*. Retrieved on May 26, 2009, from http://nces.ed.gov/nationsreportcard/tdw/analysis/2000_2001/initial_classical_descrip2000mathg12nonacc.asp
- U. S. Department of Education. (2008e). *Overall descriptive statistics for item blocks as defined after scaling, grade 12 mathematics national main assessment (accommodations permitted [R3]): 2000*. Retrieved on May 26, 2009, from http://nces.ed.gov/nationsreportcard/tdw/analysis/2000_2001/initial_classical_descrip2000mathg12.asp
- Usher, E. L. (2009). Sources of middle school students' self-efficacy in mathematics: A qualitative investigation. *American Educational Research Journal*, 46, 275-314.
- Van De Walle, J. A. (2007). *Elementary and middle school mathematics*. (6th ed.). Boston: Pearson.
- Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics*, 53, 113-138.
- Van Hiele, P. M. (1959/1984a). A child's thought and geometry. In D. Fuys, D. Geddes, & R. Tischler (Eds.), *English translation of selected writings of Dina Van Hiele-Geldof and Pierre M. Van Hiele* (pp. 243-252). Washington, D. C.: National Science Foundation. (ERIC Document Reproduction Service No. ED 287697)
- Van Hiele, P. M. (1959/1984b). English summary by Pierre Marie Van Hiele of "The problem of insight in connection with school children's insight into the subject matter of geometry." In D. Fuys, D. Geddes, & R. Tischler (Eds.), *English translation of selected writings of Dina Van Hiele-Geldof and Pierre M. Van Hiele* (pp. 237-242). Washington, D. C.: National Science Foundation. (ERIC Document

Reproduction Service No. ED 287697)

- Van Langen, A., Rekers-Mombarg, L., & Dekkers, H. (2008). Mathematics and science choice following introduction of compulsory study profiles into dutch secondary education. *British Educational Research Journal*, *34*, 733-745.
- VanLehn, K. (1980). Bugs are not enough: Empirical study of bugs, impasses, and repairs in procedural skills. *Journal of Mathematical Behavior*, *3*, 3-71.
- VanLehn, K. (1983). On the representation of procedures in repair theory. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 201-253). New York: Academic Press.
- Vogel, C. (2008). Algebra: Changing the equation. *District Administration*, *44*, 34-40.
- Von Glasersfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 3-18). Hillsdale, NJ: Lawrence Erlbaum.
- Von Minden, A. M., Walls, R. T., & Nardi, A. H. (1998). Charting the links between mathematics content and pedagogy concepts: Cartographies of cognition. *The Journal of Experimental Education*, *66*, 339-358.
- Vos Savant, M. (1990, September 9). "Ask Marilyn" column, *Parade Magazine*, p. 16.
- Wagner, S., & Parker, S. (1993). Advancing algebra. In P. S. Wilson (Ed.), *Research ideas for the classroom* (pp. 119-140). Reston, VA: National Council of Teachers of Mathematics.
- Watson, J. M., & Kelly, B. A. (2005). The winds are variable: Student intuitions about variation. *School Science & Mathematics*, *105*, 252-269.
- Watson, J. M., Kelly, B. A., Callingham, R. A., & Shaughnessy, J. M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*, *34*, 1-29.
- Watson, J. M., & Shaughnessy, J. M. (2004). Proportional reasoning: Lessons from research in data and chance. *Mathematics Teaching in the Middle School*, *10*, 104-109.
- Wei, F. Y. F., & Hendrix, K. G. (2009). Gender differences in preschool children's recall of competitive and noncompetitive computer mathematics games. *Learning, Media and Technology*, *34*, 27-43.
- Weinberg, A. (2007, October). *New perspectives on the student professor problem*. Paper presented at the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 164-170), Reno, NV.
- Welch, W. (1978). Science education in Urbenville. In R. Stake & J. Easley (Eds.), *Case studies in science education* (pp. 5-1-5-33). Urbana: University of Illinois.
- Whicker, K. M., Bol, L., & Nunnery, J. A. (1997). Cooperative learning in the secondary mathematics classroom. *Journal of Educational Research*, *91*, 42-48.
- White, C. W. (1974). A study of the ability of seventh-grade and eighth-grade students to learn basic concepts of probability and the relationship between achievement in probability and selected factors. (Doctoral dissertation, University of Pittsburgh, 1974). *Dissertation Abstracts International - A*, *35* (04), 1969.
- Whiteside, D. T. (Ed.). (1972). *The mathematical papers of Isaac Newton* (Vol. 5). London: Cambridge University Press.
- Wilkins, C. W. (1993). *Effects of using a problem solving approach to Algebra I instruction*. Mississippi State, MS: Mississippi State University. (ERIC Document Reproduction Service No. ED372934)
- Wolf, F. L. (1962). *Elements of probability and statistics*. New York: McGraw-Hill.

- Wothke, W. (1993). Nonpositive definite matrices in structural equation modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256-293). Newbury Park, CA: Sage.
- Wu, H. (2005, September). Key mathematical ideas in Grades 5-8. Paper presented at the annual meeting of the *National Council of Teachers of Mathematics*, Anaheim, CA.
- Yates, D. S., Moore, D. S., & Starnes, D. S. (2006). *The practice of statistics*. (3rd ed.). New York: W.H. Freeman.
- Zawojewski, J. S., & Shaughnessey, J. M. (2000). Mean and median: Are they really so easy? *Mathematics Teaching in the Middle School*, 5, 436-440.
- Zohar, A., & Gershikov, A. (2008). Gender and performance in mathematical tasks: Does the context make a difference? *International Journal of Science and Mathematics Education*, 6, 677-693.

DRAFT