THEY MUST BE MEDIOCRE: REPRESENTATIONS, COGNITIVE COMPLEXITY, AND PROBLEM SOLVING IN SECONDARY CALCULUS TEXTBOOKS

A Dissertation

by

CHRISTOPHER THOMAS ROMERO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Robert M. Capraro.
Committee Members,	Gerald R. Kulm
	Mary M. Capraro
	Jane F. Schielack
Head of Department,	Yeping Li

December 2012

Major Subject: Curriculum and Instruction

ABSTRACT

A small group of profit seeking publishers dominates the American textbook market and guides the learning of the majority of our nation's calculus students. The College Board's AP Calculus curriculum is a de facto national standard for this gateway course that is critically important to 21^{st} century STEM careers. A multi-representational understanding of calculus is a central pillar of the AP curriculum. This dissertation asks whether this multi-representational vision is manifest in popular calculus textbooks.

This dissertation began with a survey of all AP Calculus AB Examination free response items, 2002-2011, and found that students score worse on items characterized by numerical anchors or verbal targets. Based on previously elucidated models, a new cognitive model of five levels and six principles is developed for the purpose of calculus textbook task analysis. This model explicates complexity as a function of representational input and output. Eight popular secondary calculus textbooks were selected for study based on Amazon sales rank data. All verbally anchored mathematical tasks (n=555) from sections of those books concerning the mean value theorem and all AP Calculus AB prompts (n=226) were analyzed for cognitive complexity and representational diversity using the model.

The textbook study found that calculus textbooks underrepresented the numerical anchor and verbal target. It found that the textbooks were both explicitly and implicitly less cognitively complex than the AP test. The article suggested that textbook tasks should be less dense, avoid cognitive attenuation, move away from the stand-alone item,

juxtapose anchor representations, scaffold student solutions, incorporate previously considered overarching concepts and include more profound follow-up questions.

To date there have been no studies of calculus textbook content based on established research on cognitive learning. Given the critical role that their calculus course plays in the lives of hundreds of thousands of students annually, it is incumbent upon the College Board to establish a textbook review process at the very least in the same vain as the teacher syllabus auditing process established in recent years.

DEDICATION

To Robert and Mary Margaret who believed in me more than I believed in myself

NOMENCLATURE

20XXABY	20XX AP Calculus AB Examination Free response item Y
AAAS	American Association for the Advancement of Science
AP	Advanced Placement
FTC	Fundamental Theorem of Calculus
NCTM	National Council of Teachers of Mathematics
MVT	Mean Value Theorem
STEM	Science, Technology, Engineering, and Mathematics

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

During the 2006-2007 school year, I was assigned to teach regular and honors sections of Geometry and a section of Advanced Placement (AP) Calculus AB at a private suburban Texas high school. That fall I taught a student named Nick in one section of Honors Geometry. In that particular course we used the Larson, Boswell, and Stiff (2001) *Geometry* text. One day we were working through some mathematical tasks from the book and I disagreed with the answer given in the back of the book. The answer was absurd in the context of what was asked; the book had made a mistake. In and of itself, that answer was inconsequential. On the other hand, Nick's reaction to my correction of the book was a seminal moment in my teaching career. He confidently and succinctly stood his ground, "the book is never wrong."

Nick did not know the right answer, but my quick dismissal of the book shook his worldview. He was a young man for whom mathematics never came easily. In the absence of intrinsic confidence in his mathematical abilities, Nick had found a surrogate. He had learned to rely on the book as his mathematical Sherpa. Just as I discombobulated him, his response shook me. He was incapable of questioning the math book. His fallacious belief in the infallibility of the math textbook had handicapped his ability to learn.

Later that school year, I began my first foray into the study of textbook efficacy that would become the subject of this dissertation. The occasion was the 2007 AP

Calculus AB test. One afternoon after the administration of that test I began to hear from my students that one of the free response items was particularly difficulty. They had never seen anything like it. We used *Calculus: Concepts and Contexts* (Stewart, 2005) in my class. But apparently neither my textbook nor I had provided adequate preparation for this particular problem. I wanted to know why.

Starting in 2002, the College Board has published national means and standard deviations for each of its tests' free response items. In the summer of 2007 I learned that the national mean for problem #3 of the 2007 AP Calculus AB test (2007AB3) was 0.96 out of a possible 9. Nationally, students earned approximately 10% of the available points. Why was 2007AB3 (see Appendix A-2) so prohibitively difficult? What about that item set it apart from the 59 others posed over the past 10 years?

A quick inspection of the other scores then available revealed that the next lowest mean was 1.76 out of 9 on problem #3 from 2005AB3. Both items presented students with a table of values. Both invoked the Mean Value Theorem (MVT) and both included two solicitations for an explanation of a concept or a context among four total prompts.

Calculus students in America have different backgrounds, schools, teachers, and technology. The textbooks they share are their one pedagogical commonality. A handful of textbooks dominate the market and implicitly dictate not only ***what*** students learn but also more importantly ***how*** they learn. This dissertation is a study of the impact of textbooks on calculus learning. It is a study of the implementation of the nationwide AP Calculus curriculum in textbooks.

Literature Review

Problem Solving

The “problems” students solve are the most basic unit of classroom instruction (Arbaugh & Brown, 2004). The critical thinking, higher order reasoning, and creativity required to find a solution determine the quality of a mathematical task (Polya, 1967; Schoenfeld, 1985; Schoenfeld, 1987; Selden, Selden, Hauk, & Mason, 2000). Correlations between increased student achievement and higher quality questioning facilitated by the solving of complex problems have been documented (Lampert & Cobb, 2003). In short, questions control student learning (Manouchehri & Lapp, 2003); heightened expectation through challenging questions yields heightened achievement (Piccolo, Carter, Harbaugh, Capraro, & Capraro, 2008).

Textbooks

The dominant source of the mathematical tasks encountered by students is the textbook (Crawford & Snider, 2000; Witzel & Riccomini, 2007). American textbooks are more bloated and less focused than their international counterparts (Schmidt, McKnight, & Raizen, 1996). Teachers are largely ill equipped to differentiate between questions of varying quality (Kulm, 1994) and appropriately choose from the multitude of available tasks. It is paramount that students have the opportunity to reflect on and make connections to well-chosen high-quality problems (Porzio, 1999).

Difficulty versus Complexity

The quintessential quality of a mathematical task is its cognitive complexity (Webb, 1997), as opposed to its difficulty. The extended duration of a solution (e.g.

moving 500 boxes from one room to another) makes a task difficult not complex. There is no question about ***how*** to solve a difficult task. Alternatively, there is no immediate procedure available to solve a complex task (e.g. arranging 500 boxes to fit into a limited space). For complex tasks, the choice of mathematical tools useful to a solution is the chief element of the cognitive conundrum (Schoenfeld, 1985).

Scales of Cognitive Complexity

The Van Hiele model was developed to improve teaching by considering students' thinking (Pegg & Davey, 1998). This model, which includes five levels – visualize, analyze, generalize, deduce, and rigor, was originally developed for geometry but was purported to apply to all of mathematics (Van Hiele, 1986). The depth of knowledge scale (Webb, 1997) explicitly considered cognitive complexity (not difficulty) and has been used to evaluate the alignment between the four states' science and mathematics standards and assessment items (Webb, 1999). The four cognitive levels of Webb's scale are: recall, application, strategic thinking, and extended thinking.

Representations in Calculus

In the 1980's the calculus reform movement sought to fundamentally alter the manner in which calculus was taught (Wilson, 1997) by embracing higher quality multi-dimensional tasks that encompass a diversity of representational contexts (Demana, 1994). A national AP curriculum was developed that prescribed robust learning in terms of a rule of four – the prescription that all concepts be understood via graphical, numerical, algebraic, and verbal representations (College Board, 2012). Free response items from the AP test reflect this position; each test includes 6 free response items, each of which is a word problem involving another distinct anchor representation—an algebraic expression, graph, table of values, or combination of those. Students must equally adept at interpreting all representations.

Algebraic Representations

Algebraic representation was the singular focus of traditional calculus curriculum (Tucker, 1996). Though algebra is a powerful and essential component of mathematical learning, notational abstraction with its accompanying symbolic obfuscation often detracts from a proper understanding of the fundamental underlying concepts. Often these mathematical concepts are more clearly discernable with an alternative representation (Stigler, 1986). Calls for a multi-dimensional representational approach to learning calculus are a reflection of the raw mathematical problem solving power that accompanies versatility.

Graphical Representations

Since classroom technology became widely available, research into the role of graphical representations in classrooms has been extensive (e.g., Apinwall & Shaw, 2002; Baker, Cooley, & Trigueros, 2000; Curcio, 1987; Edens & Potter, 2008; Heid, 1988). With graphing calculators, the rote graphing of functions is trivial. When graphical representations became tractable, they became an efficient means to a conceptual end. Students' effort can be redirected towards efficiently experimenting with graphs, interpreting those graphs, and making sense of the mathematical realities reflected by the visualization.

Numerical Representations

Calculators have also made numerical representations more accessible. Tables of numerical data are available at a touch of a button and allow a student to make conjectures based on tabular patterns. Activities of this type are the most natural (and simple) for our students (Kaplan & Kaplan, 2008). Mathematicians have using tabular conjecture methods of learning through the ages. Algebraic methods dominated mathematics in the 300 years since Newton, only due to the lack of technology to facilitate efficient tabulation-based conjecture (Tucker, 1996). Children work with numerical representations before they learn algebra or how to graph an equation. Yet this representation is largely displaced in the traditional calculus curriculum (Tall, 1997).

Verbal Representations

"Without language, thinking is impossible," (Van Hiele, 1986, p.9) and without thinking, learning is impossible. Student explanations (via verbal representations) are the mechanisms through which empirical representations (algebra, graphs, and tables) are connected and explicated. Verbal discourse plays a central role in mathematics learning (Carpenter, Fennema, Franke, Levi, & Empson, 1999; Lempert & Cobb, 2003; Manouchehri & St. John, 2006). Verbal representations are the medium of mathematical learning. Without them, advanced levels of cognitive complexity incorporating rigor and abstraction are unapproachable.

Representations and Problem Solving

The essence of mathematical power is representational transfer (Goldin, 2003). Classical mathematic problems (i.e. word problems) are anchored in verbal representations and, in general, the first step towards solution is the translation of the words into some other mathematical representation (typically an algebraic expression) for interpretation. Previous attempts to create a four-level hierarchy for elementary word problems (Caldwell & Goldin, 1987) have failed to explicitly consider translations between representations. Limited attempts to explore connections between representations in calculus have been made (Porzio, 1999) but centered on student preference rather than proficiency.

Research Niche

The number of high school seniors who took at least one AP test has doubled since 2001 (Seiben, 2011) and just in the last decade the number of students taking calculus AB increased by 56% (College Board, 2012) to nearly 250,000 students each year. Calculus is a critical pedagogical juncture in the preparation of students for careers in science, technology, engineering, and mathematics (Bressoud, 2009). Despite decades of curricular reform efforts, students still entered college calculus ill prepared (St. Jarre, 2008; Selden, Selden, Hauk, & Mason, 2000).

In recent years, the College Board developed an AP course audit process (Geiser, 2008) that required teachers to acknowledge the *intended* AP curriculum (Usiskin, 1994) and, at least on paper syllabi, appropriately reflect the defined AP curricular standards (College Board, 2012). What the College Board could not guarantee, however, was the veracity of the *implemented* curriculum. What was being taught in the classroom remains unknown.

Textbooks provide a window into the implemented curriculum. They offer profound insight into both what was taught and how it was taught (Capraro, Capraro, Younes, Han, & Garner, 2012). In 1998 the American Association for the Advancement of Science (AAAS) analyzed 25 middle school mathematics textbooks and 20 middle school science textbooks, and judged only 5 satisfactory in terms of potential of promote quality mathematical learning (Roseman, Kulm, & Shuttleworth, 2001). There have been no similar studies of calculus textbooks. Without an in-depth analytic review of calculus textbooks, we cannot know whether the books are likely to help students learn (McNeely, 1997).

The roles of representations (Cunningham, 2005; Knuth, 2000; Porzio, 1999) and representational transfer (Keller & Hirsch, 1998; Romberg, Fenema, & Carpenter, 1993) in problem solving have not been adequately addressed in the literature. Representations provide a lens through which we can begin to explore the cognitive complexity of calculus textbook tasks. In this dissertation a study of a sample of mathematical tasks from popular calculus textbooks will be the first study of any kind of calculus textbooks. This study will be the first to connect representational transfer to cognitive complexity and the first to adapt the Depth of Knowledge inventory (Webb 1997; 1999) for textbook task analysis.

Article 1 (Chapter II)

The College Board's call for a multi-dimensional representational presentation of concepts is the cornerstone of AP Calculus curriculum. Students are expected to be equally adept at interpreting concepts algebraically, graphically, numerically, and verbally. Cursory inspection of the national mean scores for AP Free Response items since 2002 suggested that there was a distinct association between mean score and representational anchor/target. This article seeks to answer whether the AP calculus vision for a multi-representation learning environment in the classroom (i.e. the rule of four) has been faithfully realized?

Experience with representations plays a key role in the quality of the problem solving process (Cheng, 1999; Goldin, 2003). Algebraic representations were the primary focus of traditional mathematics pedagogies (Tucker, 1996). The danger of a single representation focus is that a preponderance of symbolic manipulation can obfuscate conceptual understanding. With the increased availability of technology, graphical representations became tractable (Edens & Potter, 2008) and offered a visual window towards student understanding. Tables of values (i.e. numeric representations) are the most basic mathematics representation but also the most overlooked in calculus (Tall, 1997). Verbal discourse is an inexplicable aspect of the mathematics learning (Manouchehri & St. John, 2006) and the foundation of every word problem.

This quantitative article attempted to verify and document the association between anchor and output representations and mean free response scores on the AP Calculus AB test in an effort to explain why students across the nation scored so poorly on Free Response Item #3 in 2007. National summary statistics were available for AP Calculus AB free response items from 2002 through 2011. Each item was a word problem built off of at least one anchor representation (algebraic, graphical, or table of numerical values). Each item included 3-4 prompts in which the students were tasked with using the representational stem to provide a solution via some distinct output representation.

The validity and reliability of the College Board test creation process, its rigorous grading processes, and the large sample sizes the existence of significant insight to be found within data in regard to representation, representation transfer, and student performance as a mirror of the current state of the nation's secondary calculus classrooms. The considered free response items (n=60) were categorized according to both input and output representations and aggregated into groups. Summary statistics for each group were computed for comparison and contrast.

Article 2 (Chapter III)

Textbooks are the single constant factor across the nation's calculus classrooms; a few publishers dominate the market for calculus books. The purpose of this article was to elucidate a model for cognitive complexity in terms of representations that would allow the analysis the mathematical tasks of common calculus textbooks. This qualitative article built on the premise that representational transfers were the essence of mathematical learning (Goldin, 2003). It surveyed the Van Hiele theory of learning (e.g. Usiskin, 1982) for geometry and recalled the bifurcation of mathematical tasks into exercises and problems (Schoenfeld, 1985). The Depth of Knowledge model (Webb, 1997) was discussed and the concepts of mechanical "difficulty" and cognitive "complexity" were distinguished.

From its historical antecedents, a new model of cognitive complexity is adapted for the purpose of textbook task analysis. The new model in explicated in five levels: Level 0 “recall,” Level 1 “apply,” Level 2 “interpret,” Level 3 “synthesize,” and Level 4 “abstract.” Each of the five levels was defined in terms of the number of representational transfers required for the solution of a given problem. Principles of Independence, Relativism, Parsimony, Inheritance, Reducibility, and Completeness were established. These principles illustrated the appropriateness and versatility of the model for mathematical task analysis. Examples mathematical tasks are provide to illustrate the cognitive intricacies of each level and principle.

Article 3 (Chapter IV)

Representational diversity is a pillar of the AP Calculus curriculum (College Board, 2012), yet students do not perform as well with free response items with either numerical anchors or verbal targets (Romero1, 2012). The College Board has instituted an auditing process of teacher syllabi to assure that courses with the designation AP meet certain curricular standards. Though the College Board provided a provided a list of sample textbooks on its website (College Board, 2012), it has not audited the contents of those calculus textbooks to assure that they meet its curricular standards.

This article investigated the following: *Does the cognitive complexity in common calculus textbooks align with the complexity expected by the AP curriculum? Is the multi-representational vision of the AP Calculus curriculum manifest in textbooks?* After adapting an instrument (Romero2, 2012) based on previous cognitive learning models (e.g. Webb, 2007), eight of the most popular textbooks were chosen for this study based on Amazon Sales Rank. Textbook lessons centering on Mean Value Theorem were selected from each book. All tasks with verbal anchors (i.e. word problems) were evaluated for cognitive complexity and representations employed and required. A parallel analysis was completed for all AP Calculus AB tasks between 2002 and 2011. Statistics for each book were aggregated into tables, along with the numbers for the AP Test items. Numbers were contrasted and results were legitimized with a subsequent qualitative analysis of a sample of tasks from both the textbooks and the AP Test.

Conclusion

Understanding mathematics is not what it used to be, the demands on our students are great and as such our textbooks must meet the challenge of empowering not only the mathematics students but also teachers of the 21st century (Usiskin, 1997). As a teacher in 2007, I trusted my textbook to guide my teaching and my students' learning. I assumed that the tasks at the end of each section would sufficiently reflect the AP curriculum I was charged to teach and adequately prepare my students for the end of the year examination. Like my student Nick, I had implicitly trusted the authors of my textbook; I, uncritically, just assumed they were right. This dissertation is about testing that assumption.

CHAPTER II

PULLING BACK THE CURTAIN ON SECONDARY CALCULUS: A REPRESENTATIONAL ANALYSIS OF AP CALCULUS AB FREE RESPONSE ITEMS, 2002-2011

During the 2006-2007 school year, I was an Advanced Placement (AP) calculus AB teacher at a private suburban Texas high school. The afternoon of the AP examination that May, I heard from my students about an “impossible” free response item they encountered on the test. It was a mathematical task for which they were ill prepared. I had failed my students and I wanted to know how and why.

Between 2002 and 2011 more than 2 million students sat for the AP Calculus AB test. (College Board, 2012). Since 2002 the College Board has released national mean scoring data on each of its free response items. The national average for this particular free response item, #3 on the 2007 AB test (2007AB3) was 0.96 points, out of a possible 9. At the other extreme of the data, the mean from #1 on the 2005 AB test (2005AB1) was 5.73 out of a possible 9. Overall 4 items fell between 1.00 and 1.99. Another 21 came between 2.00 and 2.99, 22 between 3.00 and 3.99, and 11 between 4.00 and 4.99. Of the 60 items tested between 2002 and 2011, no items other than 2007AB3 or 2005AB1 have yielded national mean scores either greater than 5.0 or less than 1.0.

The best scoring item (2005AB1) began with a graph, and asked students to find two areas and a volume. In fact 6 of the 12 best scoring free response items (i.e. with means greater than 4.0) asked students to find areas and volumes beginning with either graphs or algebraic expressions that can be quickly input into technology to yield a graph. On these items students were not required explain their reasoning; a rudimentary algorithmic application was sufficient.

The worst scoring item (2007AB3) provided a numerical table of values. Two of its prompts required students to summon intricate knowledge of calculus theorems to explain why certain facts must be true of the function represented by the table. In fact 2 of the 3 lowest scoring items (i.e. with means less than 1.77) began with a table of numerical values. Those two (2007AB3 and 2005AB3) both allowed no quick formula to be applied and implicitly invoked the Mean Value Theorem (MVT) as a means of justification for the answer to why questions.

Representations would seem play a profound role in predicting student success (or lack of thereof) on given AP Calculus AB free response items. The best scoring items were graph based and the worst scoring items were number table based. The best scoring items required simple algebraic manipulation and the worst scoring items required students to explain their understanding of a provided numerical model. Have students realized the College Board's vision of multi-representational calculus proficiency? This article investigates AP Calculus AB free response items through a representational lens. What pedagogical insights can be gleaned from the study of national mean scores on these items?

Literature Review

What Is a Representation?

External representations are systematized presentations (e.g. algebraic expressions, tables of numbers, Cartesian graphs) of information that reflect some aspect(s) of mathematical reality (Neilssen & Tomic, 1996). Learning is the process of constructing internal representations of mathematical reality (Cobb, Yackel, & Wood, 1992). External representations are the tools with which students construct internal representations (and learn mathematics). Without external representations, learning is not possible.

The emergence of representational theories of learning coincided with the global paradigm shift that has marked the dawn of the modern information age. The rise of technology in our world has presaged the introduction of cognitively guided pedagogies in our classroom. The focus of research-supported classroom best practices has shifted from the product of mathematics (i.e. answers) to the process of the construction of internal representations of mathematics.

One challenge to the establishment of cognitive based pedagogy is assessment. Traditional mechanical pedagogy is rooted in what we can easily access— answers. In the past representations were underutilized in our curriculum (Kaput, 1987). Representations allow the generation of an assessable trail of documentation that can reveal information about student thinking (Lesh, 2006). Without an explicit classroom representational focus, the assessment and implementation of cognitive based curriculum is more difficult. Representation is a key to the realization of cognitively guided pedagogy

Mathematical representations cannot be understood in isolation (Goldin & Shteingold, 2001). The degree of understanding is determined by the quantity and quality of connections between internal and external representations (Hiebert & Carpenter, 1992). The essence of mathematical power is the ability to transfer between different representations (Goldin, 2003). Partially developed internal systems of representations leave long-term cognitive obstacles and associated confidence-related barriers (Goldin & Shteingold).

Problem Solving and Representation

Though a problem solving approach to secondary mathematics learning has been advocated for years (e.g., Halmos, 1980; National Council of Teachers of Mathematics (NCTM), 1980; NCTM, 1989, NCTM, 2000), a genuine realization of problem solving in the classroom has remains elusive (Kulm, 1994; Lester & Kehle, 2003; Schoenfeld, 2004; Stacey, 2005). The defining characteristic of the problem solving process is its unscripted, non-routine nature (e.g., Polya, 1967; Schoenfeld, 1985; Selden, Selden, Hauk, & Mason, 2000). In addition to assessment challenges, the fact that problem solving is indefinable in terms of some always-applicable linear recipe contributes to its pedagogical elusiveness.

Lakatos (1976) describes the process of problem solving as following a "zigzag path of discovery (p. 42)." Adapting and refining a definition from Lester and Kehle (2003) we can precisely define problem solving as the process of resolving ambiguity via the generation of new representations from representations familiar from previous experience or knowledge. Representations are the milestones along the path of the problem solving process, based on which problem solvers zigzag—adjust their cognitive and metacognitive approaches toward success.

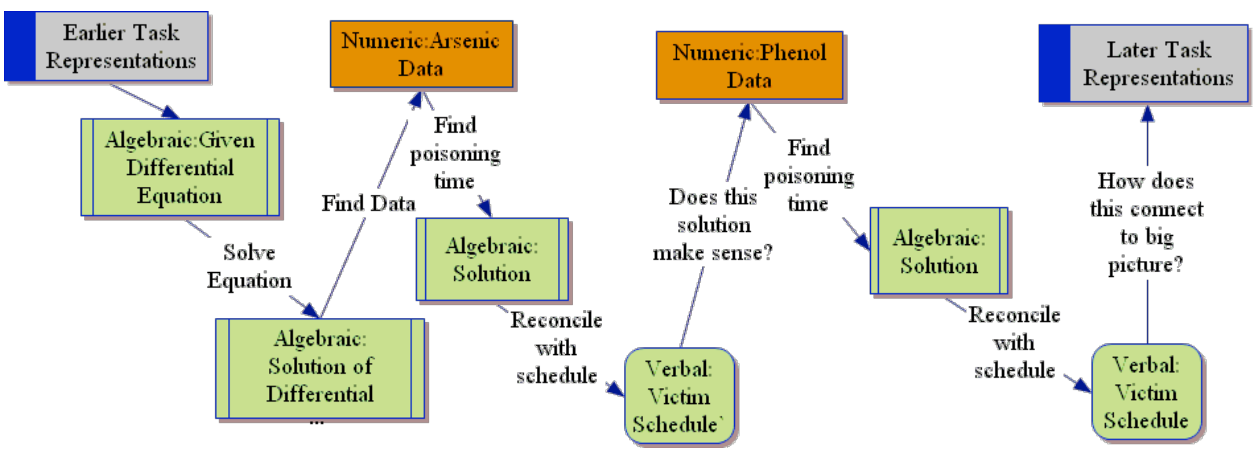
Goldin and Shteingold (2001) suggest that an avenue for overcoming the cognitive impasse with representational transfer is the explicit consideration of representational aspects of mathematical problem solving. This idea is implicitly alluded to by Polya (1973, p. 47):

> We may add auxiliary elements to the conception of our problem in order to make it fuller, more suggestive, more familiar although we scarcely know yet explicitly how we shall be able to use the elements added.

External representations are "auxiliary elements." Tables, figures, and words are not always necessary on a solution path; however, they can provide guidance when a problem solver is not yet certain which direction to follow. An examination of Polya's "dictionary" reveals that each of our familiar categorizations of representations (i.e., graphical, tabular, verbal, and algebraic) was either explicitly (e.g. draw a figure) or implicitly (e.g. mathematical induction) included as heuristics.

Adapting a diagram (see Figure 1) of Lester and Kehle (2003) as a representation of a potential problem solving process to solve a portion of a calculus murder mystery problem (Romero, 2008), we can visualize the zigzag between representations. In the portion of the task depicted, students are asked to determine which poison was used to kill the victim. The student is offered three anchor representations: an algebraic representation of a differential equation that models the decay of poison in the body, a table listing of suspect poisons with corresponding constants, and a verbal description of the victim's interactions of the previous day. The schematic in Figure 1 only depicts a portion of a possible problem solving process. Each entity in the diagram signifies a representation that a student might employ in a solution process. The solution of the murder mystery treks from algebraic to numeric to algebra to verbal. One representation leads to the next, implicitly guiding the problem solving process. Without the representations students might be left without insight into future steps.

Figure 1 Zigzag representational schematic for calculus murder mystery



Representations in Calculus

Traditionally calculus was taught exclusively via analytic representations (i.e. algebraic symbols that are manipulated with algorithms). Calculus reformers (e.g., Douglas, 1986; Solow, 1994; Tucker, 1990) noted that numeric representations (e.g. tabular data) and graphical representations (e.g. Cartesian plots), which were intractable prior to the wide-scale introduction of the graphing calculator into calculus classrooms, are underutilized in traditional classrooms. Technology made a multi-representational calculus experience possible and allowed many teachers to a take a definitive step towards the cognitively guided pedagogy advocated by reformers (Waits & Demana, 1994).

The AP calculus curriculum's *rule of four* prescribed that every calculus concept to be developed in terms of four representations –algebraic, graphical, numerical, and verbal (College Board, 2012). For example, consider the question, what does it mean to understand the concept of the derivative? An AP Calculus student is expected to understand the derivative via numerous formulas for computation (algebraic), by its definition as the limit of secant approximations (graphical), via a difference table of a discrete number of points, and contextualized as a rate of change as might be described in a newspaper article (verbal).

The Rule of Four: Algebraic

Symbolic algebraic representations were the singular focus of the traditional calculus curriculum (Tucker, 1996). American elementary textbooks have been shown to more frequently use symbolic representations to replace alternate representations and dilute the learning experience (Cai & Lester, 2005). Though algebra is a powerful and essential component of mathematical learning, notational abstraction with its accompanying symbolic obfuscation often detracts from a proper understanding of the fundamental underlying concept that is more clearly discernable via a simpler alternative representation (Stigler, 1986). This realization gave rise to calls for the inclusion of graphical, numerical, and verbal representations in the learning of calculus.

The Rule of Four: Graphical

Since classroom technology became available, research into the role of graphical representations in classrooms has been extensive (e.g., Apinwall & Shaw, 2002; Baker, Cooley, & Trigueros, 2000; Curcio, 1987; Edens & Potter, 2008; Heid, 1988). The purpose of using concrete visual representations is to mediate students' conceptual understanding of abstraction (Cai & Lester, 2005). Students' effort can be redirected towards efficiently experimenting with graphs, interpreting those graphs, and the sense-making processes necessary for conceptual learning.

Calculators are mechanisms for cognitive amplification (Grassl & Mingus, 2002); students are able to access problems that are inaccessible without technology. Calculators and their representation generation capabilities allow students can explore families of solutions to an intractable differential equation (Braiden, 2011) or climb a cognitive scaffold to understand the fundamental theorem of calculus (Schnepp & Nemirovsky, 2001). Calculators empowered the graphical representation to transcend its role as a static reflector of a correct mechanical process to a new role as one of may dynamic intermediary representations from which better understanding emanates.

The Rule of Four: Numerical

Calculators have also made numerical representations more accessible. Tables of numerical data are available at a touch of a button and allow a student to make conjectures based on tabular patterns. Kaplan & Kaplan (2008) inform us that activities of this type are the most natural (and simple) for our students. Mathematicians have using tabular conjecture methods of learning through the ages. Algebraic methods dominated mathematics in the 300 years since Newton, only due to the lack of technology to facilitate efficient tabulation-based conjecture (Tucker, 1996). Children work with numerical representations from an early age, before they learn algebra or how to graph an equation. Yet this representation is largely displaced in the traditional calculus curriculum (Tall, 1997). This fact coincides with the dearth of research into students' use of numerical representations.

The Rule of Four: Verbal

"Without language, thinking is impossible," (Van Hiele, 1986, p.9) and without thinking learning is impossible. Verbal representations are essential to the reflective cognitive action central to deep conceptual learning. (Carpenter, Fennema, Franke, Levi, & Empson, 1999; Lempert & Cobb, 2003; Manouchehri & St. John, 2006). Verbal representations (i.e., student explanations) allow student to connect physical representations (algebra, graphs, and tables) and explain "how" and "why" a solution must be correct. Modern calculus pedagogies invoke the verbal representation through contextualizing stories (Hughes-Hallett, 2006) that can make abstract mathematics concepts attainable to more students. No longer is mathematical understanding reserved to only those who are proficient with symbols.

Representational Determinism

Representations substantially affect both the effectiveness and quality of the problem solving process (Cheng 1999). Proficiency with the creation and deployment of representations towards a solution is one characteristic of expert problem solvers and a central skill requiring practice of novice problem solvers (Schoenfeld, 1985; Stylianou & Silver, 2004). The representations used to describe a function determined whether students could answer questions about it (Keller & Hirsch, 1998). That is, whichever representation is given in a problem is highly deterministic of a particular student's success with the problem.

Though a research-based underpinning of the role of representations in classroom pedagogy has begun to emerge (Keller & Hirsch, 1998), work in this area remains (Janvier, 1987; Lester & Kehle, 2003). Problem solving can be conceptualized as the process of finding a particular output representation to explain or resolve an uncertainty inherent from one or more input representations. This representational transition process is not clearly understood and should be a focus of curriculum research (Porzio, 1999; US Dept of Education, 2008).

Instrumentation

A test development committee comprised of both university and secondary experts creates the AP Calculus free response items via an established process. The validity of these items is supported by a meta-analysis of Bressoud (2009) who found that AP examination scores reasonably correlate to grades in college calculus courses. This conclusion reinforces the reliability of AP scores as a metric of future college calculus scores. The purpose of this study is to analyze available College Board data. The validity and reliability of the College Board test creation process, its rigorous grading processes, and the large sample sizes hinted at the existence of significant insight to be found within data in regard to representation and student performance as a mirror of the current state of the nation's secondary calculus classrooms.

The AP calculus test consists of two equally weighted sections of multiple choice and free response items. Multiple-choice items are not released regularly. The free response items are released each year, along with national student mean scores for those items. Each free response item includes 2 to 4 parts requiring output (or target) representations as solutions. The antecedent to each of those parts is a common anchor representation. Each stem includes a verbal representation (word problem like presentation) in combination with either an algebraic, graphical, or numerical representation as an anchor.

Each AP Calculus AB free response item is scored out of a possible 9 points. The College Board provides a scoring rubric along with each of its items. As of this writing, student statistics for 60 free response items (from 2002 to 2011) are available at the AP Central website (College Board, 2012) complete with summary statistics and scoring rubrics. Scoring data for individual parts of a given item is not available. This article seeks to answer whether the AP calculus vision for a multi-representational learning environment in the classroom (i.e. the rule of four) is evidenced by student performance on the AP Calculus AB test?

Method

For this study, each AP Calculus AB free response item offered between 2002 and 2011 was classified categorically according to given anchor representation and required target representation(s). Verbal representations were inherent in the presentation of every AP free response item and were not considered as part of the anchor portion of this study. Because it was possible for a student to generate a correct solution with different representations, the target analysis portion of this study focused on verbal representational requirements. Verbal representations requirements were identified through phrases like “explain your reasoning” or “justify your answer.”

A free response item can have as many as 3 different expected target representations included among the multiple prompts following up a given anchored problem. Each item from among the original 60 was classified based on whether its multiple prompts included elicitations for verbal representations. Of the 60 items, 27 contained at least one prompt requiring a verbal output representation. Eleven of those items required 2 verbal outputs and only one required 3 verbal outputs.

Because the national mean of an item requiring 3 verbal outputs is more representative of verbal representational proficiency than the mean of an item requiring one verbal output, individual prompts were considered distinct in this portion of the study. That is, in order to study an appropriate mean related to verbal representation proficiency, items were weighted based upon the number of verbal outputs required; a item with 2 verbal outputs was weighted twice while an item with only 1 verbal output was weighted once.

Following representational classification, 3 categorical tables of free response items were created for anchors (graphical, algebraic, and numerical) and 2 tables were created for targets (verbal and non-verbal). Each table allowed the calculations of summary statistics for items of each type for contrast. These statistics included mean, weighted mean, maximum, minimum, and number of items considered. This data then allowed for statistical analysis to compare mean student results with respect to representation with an appropriate *t*-test.

Results

The mean of the means of national student scores on AP Calculus AB free response items (n=60) from 2002 through 2011 was 3.22 with a maximum of 5.73 and a minimum of 0.96 (see Table 1). The standard deviation of the sample was 0.95. The 1st, 2nd, and 3rd quartiles were 2.52, 3.14, and 3.94 respectively. AP Calculus students scored highest on questions with a graphical anchor representation (μ=3.45) and lowest with a numerical anchor representation (μ=2.54). This corresponds respectively to the graphical and numerical anchors for the maximum (5.73) and minimum (0.96) of the entire set. It was found that weighted means (found by incorporating yearly sample sizes) were practically insignificant. All difference of weighted mean and simple mean (equal sample size assumption) was less than 0.5.

Anchor Representation Comparisons

The equality of means of items categorized by anchor representations (see Table 2) was compared with a series of independent sample *t*-tests. A statistically significant difference (α=0.05) was found when comparing the mean of numerically anchored items against the means of both the algebraically and graphically anchored items. A statistically significant (α=0.05) difference between algebraically anchored items and graphically anchored items was not found.

Table 1 Anchor Representation Summary Statistics

Anchor Representation	Max	Min	Mean	Weighted Mean	Num
Graphical	5.73	1.76	3.45	3.43	22
Algebraic	4.89	1.63	3.29	3.28	28
Numeric	3.48	0.96	2.54	2.58	11
Total	5.73	0.96	3.22	3.21	61*

Note. One item (2003 #4) included both graphic and numeric anchors

Table 2 Anchor Representation Means *t*-tests

Anchor (n)	Levene's Sig. for Eq Var	2-tailed *t*-Test Sig. No Eq Var Assumed	2-tailed *t*-Test Sig. Equal Var Assumed
Graphical (n=22) vs. Algebraic (n=28)	0.067	0.573	0.559
Algebraic (n=28) vs. Numeric (n=11)	0.470	0.011	0.011
Graphical (n=22) vs. Numeric (n=11)	0.047	0.008	0.017

Target Representation Comparisons

Over the decade between 2002 and 2011, the AP Calculus AB examinations offered students 207 prompts (see Table 3) among its 60 free response items or 3.48

prompts per item. AP Calculus students scored higher on prompts not requiring verbal outputs (μ=3.28) than they did on items requiring the verbal representation (μ=2.74). It should be noted that the number of verbal prompts has increased over the decade. From 2002 to 2005, 9 total verbal eliciting prompts were offered. From 2006 through 2008, 12 were offered; from 2009 through 2011, 19. A statistically significant (α=0.05) difference was found between the national means for items with verbal targets versus those without verbal targets (see: Table 4)

Table 3 Target Representation Summary Statistics

Target Representation	Max	Min	Weighted Mean	Num
Non-Verbal	5.73	0.96	3.28	167
Verbal	4.67	0.96	2.74	40
Total	5.73	0.96	3.18	207

Table 4 Target Representation Means *t*-test

Targets (n)	Levene's Sig. for Eq Var	2-tailed *t*-Test Sig. No Eq Var Assumed	2-tailed *t*-Test Sig. Equal Var Assumed
Verbal (n=167) vs. Non-Verbal (n=40)	0.724	0.001	0.001

Discussion

This study was inspired by the consideration of a low extreme free response item (2007AB3) and the discovery of another low extreme (2005AB3) and a high extreme

(2005AB1). The pattern of performance suggested by these extremes has, in fact, emerged following analysis. The numerical representation is the only anchor that falls below the overall mean and it is significantly different (α=0.05) from both of the other anchors. The algebraic and graphical anchors are not significantly different and both above the overall mean. There exists a gap between the student performance on numerical representation and performance with algebraic or graphical representations. Similar disparity was found between verbal and non-verbal targets. Evidence suggests that students do not understand numerical and verbal representations as well as they understand graphical and algebraic representations.

This study has confirmed the conclusion of a previous one that found anchor representations were associated with student performance (Keller & Hirsch, 1998). Student ill performance with numeric anchors likely coincides with a neglect of that representation within the implemented curriculum, just as positive student performance with graphical anchors coincides with the explicit introduction of a technology-empowered graphical representational focus into the classroom. In order for our students to perform well on tasks involving numerical representations, student need exposure to tabular based tasks in our classrooms. In fact the rule of four insists that we adopt such an approach.

Cunningham (2005) asserts that student behavior stems from instructional factors rather that cognitive ones. Textbooks dominate secondary curriculum and teachers are dependent on them as their primary source of curriculum (Crawford & Snider, 2000). It is incumbent upon textbooks, assessments, and teachers to embrace the multi-dimensional and multi-representational process of problem solving (M. Capraro, 2001; Stacey, 2005; R. Capraro & Yetkiner, 2008; NCTM, 2009).

As the chief medium through which the AP curriculum is realized, textbooks should reflect the "rule of four" focus. This article calls into question the fitness of our textbooks to facilitate the modern vision of calculus curriculum. A more in-depth study into the content of our textbooks is required. Parents, teachers, administrators, and students put faith in our textbooks and assume that they will facilitate good teaching and good learning. This research indicates that we should not uncritically accept this assumption.

Dreyfus and Eigenberg (1982) found that students with high ability preferred the graphical mode, while low ability pupils favored the tabular representation. Is this a self-fulfilling prophecy? Is a failure by textbooks to include a numerical avenue to solution an implicit indictment of the potential success of certain students? If a student who is most comfortable with tabular representations is forced to work exclusively outside of his preferred domain, perhaps he will be ill suited for success. Rote mathematical tasks are prescriptive and limiting of students' potential for success; alternative avenues via non-traditional representations to problem solving success are crucial to egalitarian student achievement.

Students' choice of representational tool is a key element of the alleviation of students' ill success in problem solving. The contrast between expert and novice performance can be, at least partially, attributed to the representations chosen to aid reasoning (Cheng, 1999). The expertise of a problem solver is grounded in the choice of representational path. Though it is instructive to study students' behavior with prescribed paths, exploration of students' behavior on tasks where he or she must choose a path offers promise of deeper insight into a more general understanding of student problem solving efficacy.

From his work with Model-Eliciting Activities Lesh (2006) suggests that a representational-tool-choice laden pedagogy offers the opportunity of success to traditionally low achieving students, while students who thrive in pedagogy centering on prescribed algebraic and graphical means may be less successful in a multi-representational environment. Our need, as teachers, for more power mathematical tasks is evident. The research (e.g. Capraro & Slough, 2008) has pointed us in the right direction with regard to more appropriate tasks. However the question remains, are powerful multi-representational tasks appropriate for the modern vision of reasoning and sense making (NCTM, 2009) available in the textbooks that dictate curriculum in our calculus classrooms?

CHAPTER III

WHAT I CANNOT CREATE I DO NOT UNDERSTAND: USING REPRESENTATIONS TO DIAGNOSE THE COGNITIVE COMPLEXITY OF MATHEMATICAL TASKS

Just prior to his death in 1988 physicist Richard Feynman scrawled the sentence "What I cannot create, I do not understand" on his blackboard (Hawking, 2001). This sentence encompasses the conundrum facing American secondary mathematics learners. I have taught calculus for 5 years—four in high schools and one at the college level. When assigned a problem to solve, my students all too often sit and helplessly wait for sudden insight that will lead to a solution. From their earliest days American students are exclusively dependent on their textbooks as their guiding light to mathematical learning. The hallmarks of powerful learning, resiliency and adaptability, are notably lacking in my students and American students in general.

The high school calculus course is the gateway between high school and college for those pursuing a Science, Technology, Engineering, and Mathematics (STEM) career. In spite of calculus reform efforts over the past twenty years, good students are still not mastering the concepts of calculus (St. Jarre, 2008; Selden, Selden, Hauk, & Mason, 2000). Previous examinations of textbooks have found that very few have potential to facilitate deep student mathematics learning (Roseman, Kulm, and Shuttleworth, 2001) and mathematics reform efforts will remain ineffective until high-quality standards-based textbooks are developed (Smith, 1994).

After an investigation of student performance on the AP examination, it was concluded in a previous article (Romero1, 2012) that the College Board's multi-representational problem solving vision that is the heart of reform tenets is not being realized in American AP Calculus classrooms. While American calculus students have different backgrounds, schools, teachers, and technology, they share one common factor – textbooks.

American textbooks rarely go beyond exposing students to the mechanics of mathematics and emphasizing the application and definitions and formulas to routine problems (Ginsburg, Leinward, Anstrom, & Pollock, 2005). The central role of textbooks in the American secondary mathematics classroom cannot be underestimated (Crawford & Snider, 2000; Kulm & Capraro, 2008; Witzel & Riccomini, 2007). They are the primary repositories of the instructional tasks that undergird student learning.

Literature Review

The selection of mathematical tasks is the most significant choice affecting student learning (Lappan & Briars, 1995). Innately mathematics is creative, practical, and functional so why do mathematical textbook tasks have the reputation of being mundane, esoteric, and ineffectual. As long as teachers are expected to "cover" textbook lessons, school mathematics will continue to be an "academic charade of procedure, form, and convention" (Gregg 1995, p. 464). Asking a student to follow a textbook step by step is the mathematics equivalent of a literature teacher teaching from a phone book (Kaplan & Kaplan, 2008).

Authentic learning occurs when students are repeatedly challenged to express and revise their current thinking methods – not because they are guided along a narrow conceptual path toward idealized solutions of their textbook or teacher (Lesh & Yoon, 2004). In other words students learn through bona fide problem solving, an aspect that is discernibly lacking in school mathematics (Lockhart,2009). Problem solving is contingent upon the generation of new representations to resolve the tension or ambiguity inherent in *unfamiliar* mathematical tasks (Lester & Kehle, 2003).

Representations

In other words mathematical representations are the objects whose creation via problem solving is alluded to by the sentence on Feynman's blackboard. The ability to interpret and synthesize given representations and develop and create new ones is a hallmark of a successful problem solver (Cai & Lester, 2005). High quality mathematical tasks require student creation of new representations, the necessary avenue to conceptual understanding (Black, Harrison, Lee, Marshall, & William, 2004).

On its website the College Board has posted national averages for each of its free response items since 2002. In 2007 students scored an average of 0.96 (out of 9) on Free Response Task #3. No free response item before or since has registered that low with students. This task provides students with a stem based around a numerical anchor (i.e. a table of data) and asked students a total of four questions: two why questions and two types of questions not commonly featured in calculus textbooks. A gap between student performances on numerically anchored problems as compared to algebraic or graphically anchored ones has previous been documented (Romero1, 2012).

Organized collections of numbers, such as a data table, are an example of the numeric type of “representations” that pervade the AP Calculus curriculum. The “Rule of Four, ” which prescribes that each calculus concept should be represented algebraically, graphically, numerically, and verbally by students, teachers, and textbooks, is a key aspect of the AP curriculum. By providing multiple concretizations of a given concept, the rule of four selectively emphasizes different aspects of a concept and facilitates a cognitive linking of concepts that facilitates a robust understanding that is more than the sum of its parts (Janvier, 1987; Kaput, 1985; Keller & Hirsch, 1998).

“Representations” are ephemeral external mirrors of reified mathematical reality (Neilseen & Tomic, 1996). Representations start as mere records of problem solving processes before becoming objects of reflection that empower deep student understanding (Lampert & Cobb, 2003). Once concepts have been internalized, a student is able to “visualize” a mathematical abstraction in his mind’s eye. He is able to draw upon his internal representations to intuit aspects of the problem solving process without needing external representations. Despite the critical role representations play in guiding, constraining, and stimulating cognition, relatively little research has considered the nature of representations in cognition (Zhang, 1997).

Once a variety of representations and their interrelations have been constructed by the student mathematics is learned powerfully (Goldin, 2003). It is representational versatility that leads to success in post calculus courses (Tall, 1992). The lack of success widely seen amongst college students (St. Jarre, 2008, Selden, Selden, Hauk, & Mason, 2000) can be attributed to a lack of multi-representational focus that would later become central reform calculus dogma. Gordon and Hughes-Hallett (1994) reported that students' appreciation of their understanding under the new representational paradigm developed. Previously students did not have to understand what they were doing to get the right answers. These students took the first step to overcoming the shortcomings of what Allendoerfer (1963) derogatorily called "cookbook calculus" by recognizing the deficits inherent in traditional one-dimensional mathematics.

The ability to transfer between representations is indicative of understanding. Few researchers have examined students' understanding of the representational connections (Knuth, 2000). The steps in students' transition from concrete representations to abstract ones are not clearly understood should be the focus of curriculum research (Department of Education, 2008). Although we do not explicitly understand the epistemological role of representations in mathematics learning, NCTM (2009) has noted that without connections among representations, reasoning and sense making facility cannot develop.

The impact of the multi-representational mathematics is vast. Students, who in the past may have failed to be successful with one learning paradigm, can escape the "one size fits all" tradition of secondary mathematics through multi-representational calculus and embrace alternative paths to success. Keller and Hirsch (1998) recalled that two dyslexic students would often misread algebraic representations but found a tractable course with graphical representations. In light of past and present students difficulties with calculus, multiple representations provide promise of a more egalitarian calculus experience and by extension more mathematicians, scientists, and engineers to lead our nation into the 22nd century.

Knuth (2000) found that more than 75% of students, when given a representational choice, elected for an algebraic solution method, even in situations when an alternative representational approach was less difficult or more efficient. In another study students preferred to use tables on the pre-test and graphs on the post-test (Keller & Hirsch, 1998). Romero (2012) found that algebra representational anchors still predominate the AP test and Tall (1997) found that traditionally the numeric representation has been missing from calculus.

Numbers are the mathematical representations that children learn first and they are the most natural representation with which to work. Yet students prefer algebra and graphs to numbers? How can this be? As the chief medium through which the AP curriculum is realized, textbooks should reflect the "rule of four" focus and equally include numerical representations among offered mathematical tasks. But do they? This study appropriates a representational lens through which we may evaluate the status of the ideals of calculus reforms as realized in our nation's textbooks and by extension our nation's calculus classrooms.

Challenging Tasks

In the 1940's Gestalt psychologist Max Wertheimer, in the work *Productive Thinking*, found that students were able to find the area of a parallelogram under circumstances they had previously seen but failed to do so when the problem was altered slightly. Schoenfeld (1988) documented students' overreliance on their textbook's examples; if they were not able to mimic algorithmic procedures they were unable to solve presented mathematical tasks. Unsurprisingly Selden, Selden, Hauk, and Mason (2000) found that student conceptual mastery is dependent on the explicit inclusion of non-algorithmic tasks as part of the curriculum.

Experiences with harder (i.e. non-routine) mathematical tasks are directly correlated to increased student achievement (Hiebert & Wearne, 1993; Knoebel, Kurtz, & Pengelley, 1994). Though various studies have analyzed textbooks (Kulm, 1999; McCrory, 2006; McNeely, 1997; Yan & Lianghuo, 2004), there are no studies in the literature that analyze calculus textbooks. Muir, Beswick, and Williamson (2008) conjecture that narrow textbook teaching approaches have implicitly encouraged students to uncritically apply algorithms to all mathematical tasks.

Students who have habitually consulted previous worked examples before attempting their own solutions have little occasion to reflect on their learning (Selden, Selden, Hauk, & Mason, 1999). Challenging instructional tasks are the key to deep learning because they provide the opportunity for reflective abstraction advocated by Piaget (Dubinsky, 1991). A student who has reflected on a number of different unfamiliar tasks will build faculty with other unfamiliar challenging tasks in general (Selden, Selden, Hauk, & Mason, 2000). As students' reflection on tasks is abstracted they become the basis of future acts of reflection in new tasks (Battista, 1999). Selden et al. (2000) found that 76% of calculus students failed to successful complete a mathematical task for which they had an adequate basis of factual knowledge. Without adequate practice with high-level tasks students will be unable to solve problems unlike those they have previously seen. High-level tasks are a necessary prerequisite for high-level achievement (Carpenter, Fennema, Franke, Levi, & Empson, 1999).

Kulm (1999) suggested that a careful analysis of content depth or complexity is required to judge whether there exists potential for students to learn mathematics with a textbook. It is this criterion that was at the center of the American Association for the Advancement of Science (AAAS) Project 2061 analysis of middle school mathematics textbooks (Kulm & Capraro, 2008). Do they provide an adequate diet of non-routine tasks capable of promoting the appropriate level of conceptual learning depth? Porzio (1997) has suggested that the calculus textbooks need to be revised and revision must be preceded by diagnostic analysis of their alignment with the multi-representational AP calculus curriculum.

The Right Questions

Because the AP calculus curriculum provides the curricular framework employed by all AP Calculus textbooks, there is little textbook variability with respect to ***what*** lesson topics are included. Every textbook will have a section that includes implicit derivatives and another will focus on related rates, for example. The central issue is ***how*** textbooks present the standard set of AP calculus topics.

In his 1997 autobiography, *The way I remember it*, eminent mathematician and analysis textbook author Walter Rudin (1997, p. 113) is quoted, “Widely used calculus books must be mediocre.” Interestingly while Rudin’s two widely acclimated real analysis textbooks are both less than 500 pages, virtually every major calculus textbook is double that size at more than 1000 pages. Unless a mathematics textbook provides students with appropriate practice, no quantity of pages or additional topical coverage will compensate for lack of mathematical task quality.

Historically, once the right questions have been asked in the right way breathtaking achievement has been accomplished in short order (Lesh et al., 2000). The first step in an assessing the appropriateness of textbook pedagogy is asking whether a textbook is asking the right questions. The right questions will start a student on a path of conceptual discovery and deeper understanding (Battista, 1999). The right questions will highlight naïve misconceptions and empower student construction of appropriate representational connections. Whether a textbook asks the right questions determines whether there exists potential for students to authentically learn mathematics with a textbook (Kulm, 1999)

The introduction of the graphing calculator into the classroom has marked the single biggest paradigm shift in the history of mathematics education. No idea, no development, no pedagogy has had a more profound ***potential*** impact on the way students learn. With the availability of a graphing calculator, some traditional textbook tasks such as the point by point graphing of a function are trivial. A more meaningful task might be to construct an algebraic function to model some plotted set of data. The first task employs a single representation and requires the student to apply a rote algorithm with little need for discussion about a trite task he would have previously completed dozens of times. The second task requires: a) the connection of three representations (algebraic, numerical, graphical), b) the recall of an appropriate parent function, c) the application of knowledge of that function when creating a graph atop the previous plot, and d) the interpretation of the appropriateness of the model. More importantly the second problem provides the opportunity for a discussion among students about the simultaneous correctness of different functional models.

The idea of what makes one mathematical task a good one and another bad is epistemological. More than fifty years ago Jerome Bruner (1960) called for the teaching for depth and continuity rather than for coverage. Over the years researchers (Knuth, 2000; Romberg, Fenema, & Carpenter, 1993) have called for studies on the developmental aspects of student learning while others have called for a focus on the content depth of textbooks (Kajander & Lovric, 2009; Kulm, 1999). No such effort has ever been completed for upper secondary mathematics (i.e. calculus and pre-calculus).

Cognitive Leveling Theories

The "depth" of a mathematical task is more properly referred to as "cognitive complexity." Historically mathematics education research has considered the notion of the complexity of mathematical tasks both informally (Keller & Hirsch, 1998; Smith & Stein, 1998) and via theoretical explications (Freudenthal, 1973). Schoenfeld (1985) distinguished between an "exercise," a mathematical situation for which a known path to a solution is available, and a "problem," a situation in which a new cognitive path towards solution must be blazed. Exercises are ineffective at revealing student's understanding or lack thereof (Lesh, Hoover, Hole, Kelly, & Post, 2000) while problems that require the highest levels of cognitive demand are positively related to the levels of student performance (Stein & Lane, 1996). In fact, curricula characterized by higher cognitive demand can help students overcome pre-existing mathematical deficiencies (Darling-Hammond, 1998; Gordon, 2006).

The cognitive leveling model of Dina and Pierre von Hiele is the most comprehensive of its kind. This theory was originally designed to improve teaching by taking into account students' thinking (Pegg & Davey, 1998). It was originally contextualized in terms of geometry and is purported by Pierre (1973) to apply to all mathematics. The van Hiele theory classified geometry understanding across five levels – recognition, analysis, order, deduction, and rigor. It would be inappropriate to apply the van Hiele theory locally, to investigate individual learning differences for instance (Pegg & Davey, 1998).

Webb's Depth of Knowledge (DOK) model employs verb analysis and contextual analysis by content experts to classify mathematical tasks into one of four levels –recall, skill, strategic thinking, or extended thinking. The DOK levels (Webb, 1997) were validated and used as one portion of a project to analyze the alignment of state standards and corresponding high-stakes tests designed to assess the attainment of those standards by students through 11th grade. This research will not employ the full program of criterion developed by Webb. Other criterion includes those considering student affective concerns and what is covered by assessment/curriculum. This study is focused exclusively on how mathematical tasks are presented.

The Model

In the subsequent paragraphs, a new diagnostic instrument of cognitive complexity is extrapolated with respect to calculus in five levels –recall, application, interpretation, synthesis, and abstraction. Unlike any previous models, this theory explicitly incorporates mathematical representations and the pathways between them. This article's penultimate section describes the principles of consistency, independence, parsimony, relativism, inheritance, reducibility, and completeness inherent in the model.

This complexity model is indebted to the work of Webb (1997, 1999) and van Hiele (Hoffer, 1983; Shaughnessy & Burger, 1985; Usiskin, 1982). Webb (personal email, 2008) acknowledged that his levels could be adapted to analyze the "complexity (not difficulty)" of textbook tasks but level definitions would have to be characterized in terms of calculus examples. The van Hiele theory has been found to be widely applicable (Usiskin, 1982) and this article seeks a global application of the two theories as a pedagogical model through which we can later analyze mathematical textbooks tasks.

This model is one of cognitive complexity of mathematical tasks; it does not consider the difficulty of tasks. Webb (1999) considered the hypothetical in which a student measures the water temperature each day for a month and then constructs a graph. This mathematical task is Level 2 according to the DOK. However if a student were to conduct a river study that considers multiple variables and constructs a model for its temperature based on those variables, the task would be classified as Level 4. The temperature measuring activity itself may be difficult (i.e. time consuming or laborious) but it is not cognitively complex (i.e. challenging conceptually). The procedure for measuring temperature is well understood. Conversely predicting the temperature of a river based on a model considering multiple variables is complex but not necessarily difficult. Completion of the second task would require blazing a new cognitive path; the first would not.

DOK analysis of how mathematical tasks are presented is related to the number of conceptual connections required and the originality of student thinking (Webb, 1997). The AP "rule of four" offers us a lens through which we can assess the complexity of calculus tasks. This model explicates a one to one relationship between the number of representations invoked and the level of cognitive complexity. In order to be deemed as aligned, mathematical tasks must elicit ideas and representational connections from students that are as demanding cognitively as the prescriptions of the appropriate curricular standards.

Level 0: Recall

The first level of the model (referred to as level 0) is a stage of recollection. In the context of mathematical questioning, items which are one-step, well defined, and algorithmic would be classified within the first level. A level 0 question requires zero representations and instead requires a single memory step such as recalling a formula. Understanding of the concept invoked is not required. "What is the formula for the third side of a triangle given two sides and the angle included between the given sides?" is a level zero item. This question can essentially be rewritten as "what is the law of cosines?" and requires recall without any use of algebraic, graphical, or numeric representations and without any understanding of what a particular definition or formula means.

I think it is important to note that particularly in mathematics, the recall of definitions is unimportant. As students move from a rudimentary understanding of mathematics to a more robust understanding they will recognize that definitions are arbitrary and relative. A student will construct (or be given) new definitions for familiar terms (e.g. parallel, normal) throughout his mathematical career. This is indicative of the cyclic process of problem solving—exploration leads to generalization leads to the establishment of a new definition that in turn leads to additional exploration.

In the context of the secondary classroom, the recall of definitions is insignificant; definitions should be given to students because they are arbitrary. In one of my calculus classes my students were confused by the definition of relative maximum as the point at which a function has a first derivative equal to zero. One precocious student pointed out that $y=x^3$ should have a relative maximum at x=0 since y' (0)=0. When looking at the graph it was apparent that $y=x^3$ did not have a relative maximum at 0. The next day I brought into class every calculus book I could find, assigned groups of students to look up the definition for relative extremes, and write the definitions on the board so that we could compare. My students were surprised that the various definitions were not the same. After discussion and subsequent exploration they discovered that a more precise definition for a relative maximum is when a function changes from increasing to decreasing.

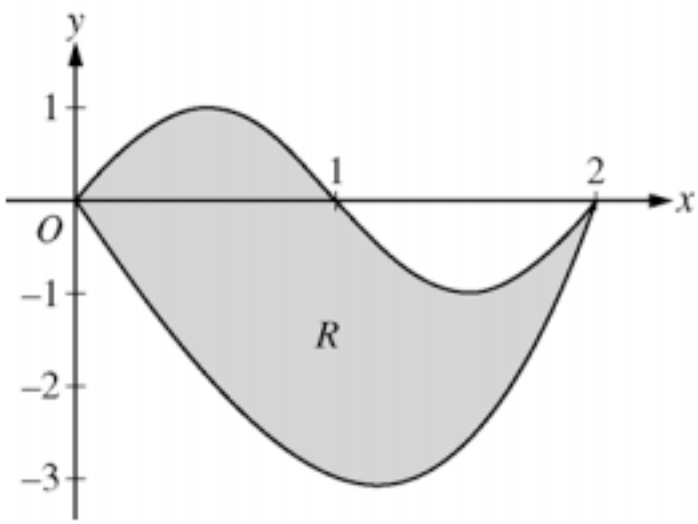
This discovery, however, was not possible without representations. Understanding of the concept of relative maximum required the use of a graph or the use of algebra; increasing/decreasing can also be understood in terms of a positive/negative first derivative test. It is incumbent on students to recognize the need for representational tools and then select an appropriate tool. In this anecdote the understanding of the definition does not require the transfer between representations but does require the selection of a single representation tool through which a solution can be achieved. This selection of a tool and the application of it towards a solution is the hallmark of mathematical tasks at the next level.

Level 1: Application

Items that require the application of rote procedures are characterized as Level I items. These items typically invoke a single representation (e.g. draw a graph of the function $y=x^2$) and hence there is no interpretation required and no need for transfer between representations. Though multiple representations may be involved such as an included graph or table, an item that does not require the use of ancillary representations is characterized as an application item.

Under van Hiele's model (Hoffer, 1983, p. 207) items at this level would require students to analyze the properties of figures such as "a rhombus has all sides equal" but they are unable to explicitly interrelate figures and properties. Under Webb's model items at this level require something more than a habitual response. In the context of calculus tasks offered across the entire nation, habitual response is difficult to qualify. What is rote for one teacher's student is not for another. Representations offer us an opening to qualify tasks independent of student context. We can classify tasks at this level via a choice of representational tool. Tasks at this level use a single representation (such as graphical for the properties of the rhombus) and require students to apply some recollection of knowledge and apply in it in a calculus context (perhaps to find a derivative). The interrelation of figures would require two distinct representations and such a task is classified in the next level.

Figure 2 Graphical anchor included with free response item 2008AB1



On part A of 2008AB1 (see Figure 2), students were asked to find the area of the region bounded by the graphs of y=sin (πx) and y = x^3–4x (College Board). A solution to this item requires the student to first algebraically or geometrically find the intersection of the two curves. This value is in turn used as a limit of integration in a standard formula to the area between the two curves. This item is Level I because its cognitive requirement is the recall and application of a cognitive tool (i.e. a rote algebraic procedure contextualized in terms of a representation). Students would be familiar with items of this type and will have practiced procedural applications of this exact type. Though this item requires two distinct iterations of an algebraic procedure, both were Level I. The aspect of the problem with the single highest cognitive complexity requirement defines the cognitive complexity required to solve the item. Though this task references a graph, this item would typically be solved entirely within the algebraic domain. The 2008 AP Calculus Free Response Solution Rubric only values algebraic methods to find the solution. This task gives students the option of finding a graphical solution (using a graphing calculator) to the algebraic equation. Even if the student were to solve the equation graphically, that part of the solution only requires 1 representation (graphical). The next part of the solution would again only require 1 representation (algebraic) and then require the algebraic solution to the integral. This item does not rise to Level 2 because those two distinct modes of representations were independent components of the solution. Without a transfer between representations a problem cannot rise to Level 2.

Level 2: Interpretation

Items that require at least 1 representational transfer and require some form of argument, explanation, or conclusion would be classified as Level 2. The central indicator involved with the distinguishing between levels 2 and 3 is representational transfer—students have to interpret one representation and output another representation. In this circumstance it is possible that the form of the representation may be the same.

For instance, a student may have to consider one graph and output another. It is important to note, however, that the two representations must be distinct.

Whereas students are mostly likely familiar with Level 1 items they encounter, items from Level 2 are unfamiliar. The van Hiele model requires the generation of solutions without reliance on rote learning (Pegg & Davey, 1998). A student will not immediately know how to solve a Level 2 item because he does not have a routine algorithm at the ready. Under Webb's DOK model, this level requires strategic thinking in which reasoning, planning, and evidence are employed towards the discovery of an original solution. It is through the reasoning of available evidence that students are able to plan a strategy that will lead to an eventual solution. This distinction between familiar and unfamiliar, between routine and non-routine are a central issue to classification and will be considered in totality after each levels' elucidation.

In the context of geometry, Hoffer (1983) noted that this level would be exemplified by tasks that require students to relate the properties of two distinct figures such as "every square is a rectangle." This recognition requires the connection of two representations— the visual representation of the rectangle and the visual representation of the square. Proofs in which the starting and ending points of the proof (e.g. Given that figure A is a square, prove that figure A is a rectangle) are provided to the student would be indicative of a task at this level according van Hiele's generalization level. Proofs that require an original conjecture followed by a justification would rise to Level 3.

In the aforementioned lowest-ever scoring AP free response stem 2007AB3 (see Appendix A-2) all four prompts would be classified as Level 2. In part a, students were asked why there must be a value *r* between 1 and 3 such that the function $h(r) = -5$ and part b asks why there must be a value *c* between 1 and 3 such that $h^{'}(c) = -5$. The requirement for an explanation in each of these prompts is distinguishing of level 2 items. These tasks are not routine because the sought after solutions are words as opposed to a standard simple numerical response (e.g. the answer is $x = 3$). Respectfully part A requires the Intermediate Value Theorem and part b requires the Mean Value Theorem. A student will have to begin with the numerical anchor (in the given table) and then translate the data within it into an appropriate verbal representation of the "why" solution. Before a student can proffer an explanation, he must examine the given numerical evidence, recall the theorems as the necessary tool, interpret the given conditions in terms of the theorem's requirements, and finally verbalize the findings in an explanation of the connection between the given conditions and the applied theorem.

In part C of 2007AB3 students are presented with the definition of a new function in terms of an integral of a previously considered function, and asked to find the derivative of this function at the point $x = 3$ (see Appendix A-2). The first step of this task is to plug 3 into the equation, a level 0 task. The next step is an application of the second fundamental theorem of calculus (FTC2) with which students should be familiar. The third step involves a representational transfer from algebraic to algebraic. What may be unfamiliar is at what point to input the 3. In one aspect $g(3)$ is a constant and naively the derivative of a constant is 0. Though the function g(x) is not explicitly defined, it is a function with a non-zero derivative at $x = 3$. Algebraic simplification, substituting in values from the tables, and calculation follow before an answer of –2 is found. This task is not a classic rote calculus application of a single concept. The implicit definition of the function (in terms of a table) is the complicating factor that differentiates this task from a Level 1 rote application. The fact that the vast majority of American calculus students had never practiced these calculus skills with a function defined by a table is the likely reason for the monumentally low score on this free response question.

In part D of AB2007A#3 students were asked to find the equation for the tangent line for the inverse function defined point-wise in the question stem. Because the task asks for a tangent line equation, the student will recall that a point and a slope will be needed to input into the point slope formula (level 0). The procedure for determining the value of the derivative of the inverse is algorithmic (level 1). These two portions are time consuming but not cognitively complex. The next step, the determination of the point is less difficult than finding a derivative but more complex. This solution step takes recognition of the reverse nature of inverse functions. The fact that $g(1) = 2$ implies $g^{-1}(2) = 1$ and $(2,1)$ is the point needed for the equation of the tangent line. This backwards-thinking step is a pre-calculus one and the most complex aspect of all 4 prompts.

In review of the prompts, an appropriate solution to the fourth prompt requires a transfer from numeric representation to algebraic representation and a segment of knowledge from a prerequisite course that very likely was not explicitly covered in the recent past. The solution to the third prompt requires an algebraic to algebraic transformation keyed by an intricate application of a theorem. An appropriate solution to the first and second prompts requires verbalizations of recalled theoretical understanding in terms of the given numeric data. This free response question invokes three major theorems— mean value, intermediate value, and the fundamental theorem, requires numerous Level 2 representational transfers, and calls for recall of non-trivial prerequisite knowledge. Proficient problem solving with such a combination of mitigating factors is exactly the goal of cognitive-based mathematics. Unfortunately students have traditionally failed to succeed with tasks of this type (Selden, Selden, Hauk, & Mason, 1999; 2000) and did not succeed with this free response item as the historically low average AP calculus score attests.

Level 3: Synthesis

Non-Routine Tasks that require several representational connections and significant planning or development would be characterized as Level 3. This level is directly akin to Webb's Level of extended thinking. Examples of tasks that might be characterized as Level 3 include the design of an experiment, the proof of a conjecture, or the connection of ideas across disciplines. This level is particularly situated towards collaboration among peers. Tasks of this type are characterized by unfamiliar circumstances. Distributed cognition efforts make verbal representations necessary, invoke the zone of proximal development, empower heightened levels of representational transfer, and allow a solution greater than the sum of its parts to emerge from a group.

This synthesis level aligns with van Hiele's deduction level in which a geometry student could independently construct and understand a proof in terms of a sequential application of theorems and postulates. An original proof requires planning and development indicative of this level. It is unlikely that such a proof would be completely without false starts and appropriate course correction. A student would be able to make conjectures about interesting mathematics and work towards a resolution to his conjecture based on his understanding of the mathematics system. The student would be required to complete multiple representational transfers, constructing an interdependent sequence of logical statements corresponding to algebraic and graphical representations. The verbal mode of representation is particular pertinent as the medium through which a student constructs his proof.

Capraro and Slough (2009) cite an exemplar of a calculus project that would be classified as a Level 3 task. In the "Who Killed Bob Krusty?" exemplar, students are provided with an extended word problem, a Clue-inspired scenario in which a murderer must be discovered based on three sub-tasks. In the first, students must naively apply (level 1) a differential form of Newton's kinematic laws. In the second, students must apply Newton's law of cooling in a novel manner (level 2) to determine time of death. In the third, students must deduce (level 3) which suspects were viable based upon time of death, a given differential equation, and a table of physiological decay rates for various poisons. Because the highest cognitive complexity contained within this task is level 3, the Bob Krusty problem is level 3.

Who Killed Bob Krusty has a simple solution; the solution required is an essay justification of reasoning. The solution requires the transfer among numerical, algebraic, and verbal representations. Though graphical representations are not required it is possible a student might employ them. The solution requires students to connect a problem that is more typical of a science classroom than mathematics. Students must connect reasoning from different disciplines and must synthesize three distinct clues discerned via calculus reasoning from the given verbal elucidation. There is no single path to a solution and this problem falls short of the final level requiring abstraction.

Level 4: Abstraction

Because of the extended planning and time requirements, items of this complexity are infeasible on the AP calculus examination but are feasible in textbooks. One might attribute a lack of this complexity to the all-for-the-test mentality that pervades modern mathematics classrooms. Thornton (1998) noted that in geometry textbooks the dearth of tasks at this level corresponded to the lack of tasks of this complexity in the classroom. Classrooms lack complex tasks because books lack them. Books lack them because the AP test lacks them and student learning is sacrificed. The tail wags the dog and students, teachers, and textbook authors are locked in a vicious cycle of good enough to pass but nothing more.

Van Hiele posited an ultimate stage where students could abstract their understanding to an extent such that he or she could compare distinct axiomatic systems in the absence of a concrete model. Later in his career van Hiele recognized the “rigor” level is inconsequential to the secondary classroom experience (Pegg & Davey, 1998). This realm is reserved for the complexity and depth characteristic of graduate level, proof based mathematics. In comparison Webb’s DOK lacks a “rigor” or “abstraction” stage but as van Hiele realized such a level is impractical.

For example Andrew Wiles’ proof of Fermat’s Last theorem, perhaps history’s best known unsolved mathematical problem, would be classified as a level 4 task. With a proof of more than 100 pages, one can only imagine the number of representational transfers Wiles' solution encompasses. The solution requires immense mathematical creativity to connect a series of seemingly unrelated conjectures. The solution is definitely abstract and makes use of new abstract branches of mathematics (e.g. group theory) that do not have concrete models and had not yet been developed when the problem was formulated.

For all practical purposes, this level will only characterize an extremely small number of tasks found in the high school curriculum. A problem of this level would not appear on the AP calculus test due to extreme time constraints. It is unlikely that any item of this level would appear in a mathematics textbook due to the inherent ambiguity of the solutions of such items. Due to the nature of these problems at this level, there are no short, neat solutions and they cannot be encapsulated in a solutions manual. Level 4 tasks are open-ended and characterized by ambiguous solutions. These solutions could take multiple forms. Two absolutely correct solutions could be completely dissimilar, if a solution exists at all.

The Instrument

Problem Solving proficiency has been the defining characteristic of mathematics knowledge. Schoenfeld (1985) defined the distinction between problems and exercises. That finding is mirrored in the differentiation of exercises as levels 0 or 1 and of problems as levels 2, 3, or 4. Selden et al. (2000) summarized the bulk of instructional task research in textbooks by noting that the vast majority of tasks are excises not problems.

The van Hiele's were interested in ways to develop insight, the ability to perform unfamiliar mathematical tasks (Hoffer, 1983). A student cannot solve the same problem twice. A solution to a problem must be original and unfamiliar. Novel insight is developed via understanding what, why, and when to apply knowledge in order to solve problems (Hoffer, 1983). The greatest naïve misconception of upper secondary mathematics is based in the "what—how" dichotomy. Students do not do poorly because they lack knowledge; too often teachers try to teach mathematics as a discrete catalog of skills. Students perform poorly because they lack the ability to apply their knowledge. The main goal of a mathematics teacher is to empower students to learn *how* to think not just *what* to think.

What remains is an instrument for the diagnostic analysis of instructional tasks. That instrument employs the cognitive complexity definitions as the basis of its indicators and is included as appendix C. Each level is based in part of the theoretical foundations laid by van Hiele, Webb, and their adherents who have further explicated those theories. Because this instrument was created on the "shoulders of giants," the dogmatic aspects of its predecessor theories are inherited by this model and allow the following principles to be established as representative of this effort. These principles when considered in combination with the discretely defined cognitive complexity levels allow a teacher or researcher to effectively analyze the potential of student learning through the instructional tasks found in calculus textbooks.

Principles

Independence

Consider the calculus task $\int_0^1 \frac{x}{x^2 - 9}\,dx$. This problem can be solved by one of at least three algebraic methods: u-substitution, trig substitution, and partial fractions. It could be solved approximately with a graphical/numerical approach using Simpson's Rule, the Trapezoid Rule, or Riemann sums. The choice of procedure is irrelevant to cognitive complexity. This task is a level 1 item; a task's cognitive complexity is ***independent*** of approach used to solve it. In this case, a u-substitution approach is rather straightforward. The partial fraction approach is a bit more algebraically dense and the trig substitution is more daunting still. In effect the approaches require prerequisite procedural skills from Algebra I, Algebra II, and pre-calculus respectively. The choice of tool will change the difficulty of the item, however the length of time or amount of effort required does not affect the complexity of the item. The cognitive hurdles one must overcome to find a solution only affect the complexity.

Relativism

Because a student cannot ever solve the same problem twice (Selden et al., 2000), the cognitive complexity of an instructional task is relative to the student. The aforementioned integral would be classified at one level for a calculus student because he has had the prerequisite precalculus course and at a higher level for an algebra II student. Partial fractions, u-substitution, and trig substitution would not be known to an Algebra I student. He has not yet "discovered" or "learned" that content. To solve the integral, a calculus student would have to choose his tool from a cognitive toolbox. An Algebra I student would be required to develop those tools before he could attempt to solve the task. Because impromptu learning is required for the Algebra student, the task would be classified at a higher level than it would be for a calculus student.

Parsimony

The graphical/numerical approach to a solution would require applying a geometric algorithm to a graph of the integral's function argument and then translating that approach into a calculation. A researcher might be tempted to code this item as Level 2 because it *could* require a translation from a graphical representation to an algebraic one. However, each item must be coded according to the cognitive steps that *must* be used. Students, teachers, and textbooks prefer the most common (i.e. most direct) route to a solution. It is this most common cognitive requirement that determines the cognitive complexity.

Inheritance

A mathematical task of a given level will encompass all levels that come before it in the model. That is, level 2 tasks will inevitably include sub-steps, which are level 0 and level 1. Every mathematical task will unavoidably involve some recall step; good mathematical tasks are grounded in basics that are established and well known to the student.

Completeness

This model of cognitive complexity is complete; every mathematical task can be classified at some level on the continuum. Both the most elementary counting task learned by toddlers and the most complex unsolved problems that baffle the smartest of mathematicians can be classified.

Reducibility

A task is cognitively leveled based on students' thinking. The context of the statement of the task can undercut the intended cognitive complexity of that task. 2007AB3a asked "why there must be a value *r* between 1 and 3 such that the function *h(r)=-5"* and was classified as level II. However if the question were rewritten "Using the intermediate value theorem, why must there be a value *r* between 1 and 3 such that the function *h(r)=-5"* it would be a level I task. The inclusion of the introductory phrase prescribing the solution route removes the cognitive step of choosing a mathematical tool and hence reduces the cognitive complexity of the item to application.

Implications for Teaching and Research

Previously I taught mathematics at a large public high school in an affluent suburb of Houston, Texas. Cindy is a veteran teacher of twenty years and my department head. Each year her principle concerns are students' passing the April administration of the mathematics portion of the Texas Assessment of Knowledge and Skills (TAKS). At a February department meeting Cindy expressed frustration with our planning for last minute TAKS preparation, a series of pull-outs and mandated tutorials covering troublesome topics from the full scale TAKS practice test the entire school took in January. In the course of this meeting Cindy was upset by the endless cycle of reteaching, in particular that of slope. She acknowledged that the 11th grade students “had been taught slope 5 or 6 times and they knew it for the test but still can’t do it on TAKS”. We were going to try to teach it to them one more time.

Cindy’s posture and affective state did not suggest any confidence in possible success of our eleventh hour efforts. She was resigned to failure once more and was clueless as to what we could do differently to affect a better outcome. We are going to try to teach them a 6th or 7th time because maybe this time it will stick or if not at least they might forget right after the TAKS test. They may not actually learn the material but hopefully our school will earn the ballyhooed Texas Education Agency Exemplary rating.

The cognitive complexity diagnostic instrument developed in this article is the tool that every school needs. It is this tool that can provide teachers with a pedagogical road map to successful student learning. The basis of this instrument (particularly within a geometry context) has been available to teachers for decades but van Hiele is nowhere to be found in my suburban Houston high school. Our teachers ascribe to an old fashioned rugged individualistic and blind pedagogy. Just as in so many other schools, my school idiot proofs its curriculum. We follow the textbook for better or worse. Whatever happens we trust the mathematics authority of the textbook. Ron Larson may have failed us the first six times, but we continue to trust him year after year to guide our students’ learning experience.

Five years ago I taught an honors geometry section using Larson’s geometry and my students struggled mightily with proofs. In the book the author presents tasks in which students are asked to recall theorems and postulates (level 0) and then fewer tasks in which they are interpret this properties in very specific contexts (level 1). The author asks the students to complete proofs in which students logically traverse from one statement through 3, 4, or 5 intermediary steps to another (level 3). My students could do level 0 and level 1 tasks but could not do level 3 tasks. I recognized the missing level 2 and recalled the continuous nature of progression through the levels. My students could not be successful at level 3 tasks without first being successful at level 2 tasks.

In this circumstance, the van Hiele model provided me with a prescription for intervention. I printed out 10 proofs (even more challenging than those offered by the textbook's authors), cut the statements out without any numbering, and required my students to work in groups to reconstruct the proof. In effect what I did was I took level 3 tasks and made them level 2 tasks. I asked my students to relate properties and figures because the textbook had failed to consider this level. I encouraged my students to make the level 2 connections first so that they could ascend to the level 3 connections later.

My geometry textbook was undercutting my students' learning. Without a van Hiele inspired intervention, I suspect my students learning would have been significantly retarded. It is my best guess that the multiple unsuccessful attempts to teach slope alluded to by my department head are directly attributable to textbooks that are not properly aligned to the best cognitive practices known to research. The van Hiele model allowed me to help my student cross the necessary cognitive bridge. This instrument elucidated by this article promises to do the same for other teachers and their students.

Although the geometry context for cognitive complexity has been the most fully explicated, there remains a dearth of research on the cognitive complexity of the textbook learning associated with other courses antecedent to calculus. In particular Algebra is rife with need for complexity analysis. This research has filled part of the gap with respect to calculus but still our education system is in desperate need of a large-scale textbook analysis effort for geometry, algebra II, precalculus, and calculus.

We can use a stage metaphor to highlight the inherent curricular issues in classrooms. The student is akin to an actor, the textbook to the script and the teacher to the director. It is expected that the curriculum will be rehearsed by student-actors as directed by their teacher-director to be later performed on a test. The script consists of the catalog of exercises-lines a student is expected to master-perform. For fear of not meeting standardized expectations the administrator-producers cannot allow deviations from the script.

The only way in which our classrooms can improve is through thoughtful textbook refinements. This cognitive complexity diagnostic instrument is the tool we need to begin the process through which we can attempt to break the textbooks' firm entanglement with stale, sterile, and unyielding curricula. It is naïve to believe we can break the textbooks' hold on our curricula. Maybe this instrument is the tool we need to align our textbooks with challenging, multi-representational tasks indicative of modern, enlightened, cognitively guided mathematics learning.

CHAPTER IV

THEY MUST BE MEDIOCRE: A COGNITIVE ANALYSIS OF POPULAR SECONDARY CALCULUS TEXTBOOKS

Walter Rudin, the author of the compact 325-page classic text *Principles of Mathematical Analysis* asserted, “Widely used calculus books must be mediocre” (1997, p. 113). I used “Baby Rudin,” as it is affectionately known, for grad school and analysis class and remember the work for its brevity, succinctness, and raw power. It is the tool that thousands of young mathematicians have used for the last half-century.

This research began on a spring day in 2007 when my calculus students took the Advanced Placement (AP) Calculus AB test. After the test my students spoke of a free response item that was “impossible.” Little did I know at the time, but that particular problem was more impossible than any AP Calculus AB Free Response had ever been. College Board, the maker of the AP test has released scoring data on every one of its free response items since 2002. In more than at decade of problems, more than 60 in total, no other item has yielded a lower national mean score of 0.96 out of 9 (College Board).

More than 211,000 students sat for the AP Calculus AB test in 2007. Calculus students in America have different backgrounds, schools, teachers, and technology. The textbooks that they share are their one pedagogical commonality. A handful of textbooks dominate the market and implicitly dictate not only ***what*** students learn but also more importantly ***how*** they learn. Without an in-depth analytic review of calculus textbooks, we cannot know whether those books are likely to help students learn what they need to learn (McNeely, 1997).

This article is the culmination of my action research to explain why this particular problem was so egregious. This article is framed by a statistical curiosity on the AP Calculus AB Exam and the indubitable truth of the importance of representation in learning calculus. It was shaped by the fundamental and elusive role of the mean value theorem, and explores the role of textbooks as the most impactful curriculum influence in secondary AP Calculus classes across the nation. It is a quest to find a prescription that might condition a better higher secondary mathematics curriculum future.

Literature Review

The transition from high school to college mathematics (a role which calculus now occupies) is a most critical juncture in the preparation of individuals to meet the mathematical demands for STEM fields for the 21st century (Bressoud, 2009). Although, it is claimed that reform efforts have moved calculus from a course of meaningless symbolic manipulation to a "lean and lively" course which, through multiple contexts, affords its students the opportunity answer to why questions (Ferrini-Munday & Graham, 1991) ill reports of student learning persists (St. Jarre, 2008; Selden, Selden, Hauk, & Mason, 1999; 2000). Previous research studying student results on the annual Advanced Placement (AP) Calculus examination (Romero1, 2012) found patterns of student difficulties with problems anchored with numerical representations or targeting responses with verbal representations.

In *Foundations for Success*, the US department of Education (2008) stated “the system that translates mathematical knowledge into value and ability for the next generation is broken and must be fixed” (p. 11). An improvement in our system of education requires an improvement in our textbooks (Crawford & Snider, 2000; Kulm, 2001; Schmidt, 1996). Current textbooks are often divorced from research-supported best practices and are instead a reflection of the antiquated pedagogy that is all too comfortable to entrenched educators and promulgated by profit-seeking publishers (Battista & Clements, 2000; Schwartz, 2006). To understand the impact of school mathematics, research must consider the curriculum that is enacted via textbook tasks (Smith & Star, 2007). Textbooks are the most common element across American classrooms and are de facto national standards (Schmidt, McKnight, & Raizen, 1997) for the providence of mathematical instructional tasks. Textbooks are the fossil records of American pedagogy (Capraro, Yetkiner, Ozel, Capraro, Ye, & Kim, 2009). Teachers decide what to teach, how to teach it, and assign instructional practice based largely on what is contained in the textbook (Reys, Reys, Tarr, & Chavez, 2006).

The question of what is expected from good mathematics textbooks has been one of the most contentious questions of decades past. In its *Principles and Standards* (1989, 2000) the National Council of Teachers of Mathematics (NCTM) tells us that “students must learn mathematics with understanding” and become “…flexible and resourceful” problem solvers by engaging in complex mathematical tasks. Mathematical power is not mechanical proficiency or general conceptual understanding or test taking success. The functional unit of mathematical power is cognitive proficiency demonstrated through problem solving facilitated by creative critical thinking. There is a correlation between higher level questioning and increased student achievement (Lampert & Cobb, 2003).

Cognitive theories of learning hold that knowledge is constructed and such construction is facilitated by social contexts. Challenge is the keystone and was the essence of Cobb's (1988) call for teachers to affect a dialectical learning environment which can allow students to transcend the basic stage of understanding and move towards the more encompassing metacognitive stage, a revision of the Pigetian notion of reflective abstraction (Pandiscio & Orton, 1998).

It was Vygotsky's work, introduced to the U.S. during the 1960's, that leads us to the realization that all higher order cognitive skills originate in and develop via the internalization of individuals' interactions with one another (Hung & Chen, 2001; Vygotsky & Kozulin, 1986). A challenging question requires equivalent dissonant representations of a mathematical situation. That dissonant situation, in turn, is the precursor to the struggle to solve a problem, which is abated by the social negotiation inherent in problem solving pedagogy. It is the challenging question that leads to dissonant representations that are reconciled via communicative struggle among peers that maximizes a student's opportunity to internally construct mathematical understanding and authentically learn (Goldin & Shteingold, 2001).

The basic dichotomy between problems and exercises is inherently predictive of the potential of mathematical tasks to promote good learning (Schoenfeld, 1985). No person would find the task of lifting a 200-pound suitcase complex. Everyone knows how to lift a suitcase. This task is a cognitive “exercise” because it is commonly understood how to lift a suitcase no matter how arduous it might be for a person who is not physically strong (Webb, 1997). Conversely, a task such as finding the set of possible integer angles in any triangle with integer side lengths is not particularly difficult. It is complex. Brute force might work for the suitcase but guessing and checking for triangle task would be intractably difficult. For a smart problem solver, however, the solution is quick if he recognizes that the law of cosines (Scher & Goldenberg, 2001) is an appropriate cognitive tool and can intuit that 60 degrees, 120 degrees, and 90 degrees are the only integer angles whose cosines are rational (either ½ or 1). Almost every student can learn by finding a solution to an unfamiliar problem (Selden, Selden, Hauk & Mason, 1999; 2000). Almost no student will learn by completing a mundane task.

In 1957, Dina and Pierre van Hiele elucidated a five-stage model of cognitive understanding applicable specifically to geometry. Therefore, cognitive models should serve as the foundation to inform our understanding of student development and as a likely beginning for examining the effectiveness of our textbooks in fostering student learning. For instance, Hoffer (1983) noted the tendency of traditional geometry texts to begin at van Hiele level 4, ignoring the natural scaffold process to deduction. In *Criteria for Alignment* (1997), Norman Webb elucidated a depth of knowledge (DOK) scale for evaluating mathematics assessment items. Whereas,the van Hiele model was intended as a developmental model directly applicable to a specific student, the DOK scale was intended as a diagnostic tool applicable to mathematical tasks approachable by some hypothetical student.

Staging theories of cognitive development, such as Bloom's Taxonomy, are directly antecedent to staging theories of learning, such as van Hiele's, which by extension provide a basis on which complexity analysis, such as Webb's, can be adapted (Webb, 2007) via the problem solving theories of Schoenfeld to analyze mathematics textbook tasks. A new model, complete with indicators, was developed (Romero2, 2012). This model provides a single conceptual framework elucidated in terms of calculus examples. A graphical depiction of the model is provided in Appendix C.

Previous studies have focused on lower secondary textbooks and found none to be high quality (Roseman, Kulm, & Shuttleworth, 2001; Capraro & Kulm, 2008) while others (e.g. Battista & Clements, 2000) suggest that current commercial textbooks are doing more harm than good, there have been no studies on Calculus textbooks. The College Board, the company that creates the AP Calculus test, provides a list of 21 "example" textbooks (College Board) that "meet the curricular requirements" but disclaims the list as, "not exhaustive and the texts listed should not be regarded as endorsed, authorized, recommended, or approved by the College Board." Without a calculus textbook review, we cannot know whether its material will help students learn as expected (Mullis, 1996).

A central tenet of AP Calculus (College Board, 2012) is that the curriculum should ubiquitously embrace a multiplicity of representation; each concept should be presented: numerically, graphically, verbally, and algebraically. This policy is a reflection of the underlying principle of the calculus reform movement (Stanley, 2002) and recognition of the indisputable truth that the ability to translate between representations is the essence of individual mathematical power (Goldin, 2003). Previous research (Romero1, 2012) found patterns connecting representations and student performance on historically low-scoring AP calculus free response items. Specifically students underperformed on free response items featuring numerical stems or requiring verbal responses. Conversely students performed better on questions with algebraic and graphical stems. Obviously, there is a question as to whether the textbook enactments meet the expectations tested on the AP exam. That is, the intended curriculum may not align the textbooks' diversity of representation.

Understanding mathematics is not what it used to be. The learning demands on our students are great and our textbooks must meet the challenge of empowering not only the math students but also the math teacher of the 21st century (Usiskin, 1997). The viability of the problem solving content of calculus textbooks is paramount to the issue of whether students will be able to succeed. A teacher may be doing an excellent job of teaching problem solving inside his classroom, but if practice tasks available in his book are insufficient or inadequate, his students will likely be unable to reify the higher order problem solving skills required by the AP curriculum. Without problems of an appropriate cognitive complexity or a presentation consistent with modern understanding of how students think and learn, it may be impossible for a good teacher to effectively teach a calculus student in the 21st century.

Research Questions

Although various issues regarding student learning in calculus (Aspinwall & Shaw, 2002; Hughes-Hallett, 2006; Selden, Selden, Hauk, & Mason, 1999; 2000; Tall, 1997) and issues regarding textbook content (Caparo, Yetkiner, Ozel, Capraro, Ye, & Kim, 2009; Ginsburg, Leinwand, Anstrom, & Pollock, 2005; Kajander & Lovric, 2009; Kulm & Caparo, 2008; Mullis, 1996) have been explored over the years, there is no extant research concerning the role of textbooks in calculus learning. Using a previously developed instrument (Romero2, 2012) this study investigated the following: *Does the cognitive complexity in common calculus textbooks align with the complexity expected by the AP curriculum? Is the multi-representational vision of the AP Calculus curriculum manifest in textbooks?*

Method

The subjects of this study were the mathematical tasks of best selling commercially produced AP calculus textbooks. From AP calculus AB textbook list (College Board), the eight best selling textbooks (see Table 5) were selected. To determine the best selling textbooks Amazon's sales rank number was used; the higher the rank, the more popular the book. It should be noted that in some cases, the College Board included multiple versions of texts by a given author in its list. Because authors offer multiple synoptic editions of their textbooks with subtle differences, this study only considered one book per author team regardless of Amazon's ranking.

From each textbook, a sample of two lessons, covering the concepts of mean value theorem and average value of a function, were selected. These two lesson topics were chosen because they were the chief conceptual subject matter of the historically low scoring AP Calculus AB Free Response Item (Romero1, 2012) that inspired this study.

After the index of each of the 8-subject textbooks was searched for "Mean Value Theorem" and "Average Value," the appropriate page numbers were noted. From these pages the sections (see Table 6) in which these two concepts were first developed was noted. Prior to the coding of the tasks from any textbook lesson, all AP Calculus AB free response items since 2002 with available summary statistics, were coded using the Cognitive Complexity model (see Appendix C) previously developed (Romero2, 2012). These data served as a control for comparisons after the coding of the 15 of the 16 textbook lessons. Thomas' text did not cover Average Value.

Table 5 Sample Textbooks by Amazon Sales Rank

First Author	Abbreviation	Amazon Ranking #	Publisher	Edition; Year
Larson	LA	28,727	Houghton Mifflin	8; 2006
Rogawski	RO	30,726	W.H. Freeman	1; 2008
Stewart	ST	43,888	Brooks/Cole	3; 2005
Finney	FI	46,583	Pearson	3; 2007
Anton	AN	52,092	John Wiley	8; 2005
Thomas	TH	68,598	Addison Welsey	11; 2007
Foerster	FO	106,814	Key Curriculum	2; 2005
Hughes-Hallett	HH	128,456	John Wiley	4; 2005

Every mathematical task solution opportunity offered to students within each of the 15 lessons was coded using the Cognitive Complexity model.. A previous study (Romero1, 2012) found that all AP Calculus Free Response Items were word problems (i.e. they were characterized by a verbal anchor). Because of this fact, all textbook tasks not including a verbal anchor were not considered further. For each task with a verbal anchor (i.e. word problem), alternate anchor representations, required output representations and cognitive complexity were noted. All coded data were compiled into summary tables (see Tables 7 & 8) for comparison and contrast.

In instances where a given item asked more than one question, the question with the highest complexity requirement dictated the complexity for the whole item. In instances in which multiple representational inputs or outputs were encompassed, multiple representations were noted on the coding documents. Therefore, while the total number of items in the complexity column summed to the total number of items, it was possible for the total in the representation columns to not equal the total number of items. Additionally, some tasks had no representational anchor beyond the verbal representation inherent in word problems.

Validity for this quantitative analysis was based on the work of Webb (1997, 1999, 2007) whose Depth of Knowledge inventory was antecedent to the Cognitive Complexity Model (Romero2, 2012). Because this research was an individual effort, inter-rater reliability was not a concern. However, intra-rater reliability, with respect to consistency of the single rater, was considered and steps were taken to minimize rater drift over the duration of the coding process. After coding all the lessons from the 8 textbooks, the first one-third of the lessons was recoded to control rater drift and a dependability of 96% was established.

Table 6 Sample Lessons

Section	Lesson	Book	Section Title	Numbered Items	Page
3.10	MVT	HH	Theorems about Differentiable Functions	28	155
4.3	MVT	ST	Derivatives and the Shapes of Curves	58	279
4.3	MVT	RO	The Mean Value Theorem and Monotonicity	71	230
4.2	MVT	FI	The Mean Value Theorem	63	196
5.7	MVT	AN	Rolle’s Theorem; Mean Value Theorem	43	330
3.2	MVT	LA	Rolle’s Theorem and the Mean Value Theorem	85	174
5.5	MVT	FO	Mean Value Theorem and Rolle’s Theorem	41	211
4.2	MVT	TH	The Mean Value Theorem	68	277
5.3	AVF	HH	Fundamental Theorem and Interpretations	37	259
6.4	AVF	ST	Average Value of a Function	20	467
6.6	AVF	RO	Setting up Integrals: Volume, Density, Average Value	60*	387
5.3	AVF	FI	Definite Integrals and Anti-Derivatives	53	287
7.6	AVF	AN	Average Value of a Function and its applications	32	476
4.4	AVF	LA	Fundamental Theorem of Calculus	106	286
10.3	AVF	FO	Average Value Problems in Motion and Elsewhere	20	509

Following the aggregation into metacategories of the textbook and AP Test tasks, post hoc internal reconciliation yielded a series of reflections regarding the comparison of textbook tasks with AP test tasks. The Quantitative results of this study were legitimized with a qualitative assessment (Johnson & Onwuegbuzie, 2004). A selection of 16 textbook tasks and 12 AP test tasks (see Appendix A) were chosen to demonstrate how textbook tasks demonstrated significant cognitive discrepancies in comparison to the test items.

Results

The composite results of the analysis of representation and cognitive complexity in the sample of free response items found in the most popular calculus textbooks was given in Tables 3 and 4, respectively. This study considered a total of 555 mathematical textbook task prompts sampled from sections on mean value theorem and average value of a function and contrasted the contexts and expectations of those prompts with 226 prompts found on the AB Calculus AB test from 2002 through 2012. One important side note was that textbook tasks without a verbal anchor were not considered. Thus this study ignored the “drill and kill” exercises that were not assessed on the AP Calculus Test.

Representations

Some of the lowest scoring tasks from the AP Calculus AB test over the past 10 years centered about numerical anchors (i.e. tables of values). These low scores nationally can undoubtedly be attributed to the lack of numerically anchored tasks in calculus textbooks. Whereas, more than 20% of AP Calculus prompts were anchored by a table, 5 of the 8 textbooks considered in this study featured less than 2% of its tasks anchored by tables. Given that teachers choose which tasks to assign, it was likely that many students had never seen a task with a numerical anchor invoking the mean value theorem before the test. Likewise there was an underrepresentation of graphical anchors and verbal targets. More than 39% of AP Calculus prompts were anchored with graphical representations while less than 25% of textbook tasks were. More than 29% of AP Calculus tasks required an explanation or the answer to some why question. Less than 20% of textbook tasks offered a similar requirement.

The sample textbooks included too few opportunities for students to develop understanding of calculus concepts through interpretation of graphical and numerical anchors. The books also included too few opportunities for student to answer the "why" questions and further develop understanding through verbalization. Based upon the results, secondary calculus textbooks lack the diversity of representational anchors found on the AP test. It must be concluded that the textbooks do not adequately realize the AP Calculus curriculum's rule of four (College Board, 2012).

Cognitive Complexity

On the 2007 AP Test Free Response Item #3 (College Board, 2012), students were twice required to demonstrate verbally their understanding of the application of the mean value theorem and the intermediate value theorem, respectively. Each instance was worth 2 points. These prompts would fall on Level 2, "Interpret" on the Cognitive Complexity model previously elucidated (Romero2, 2012). The remaining 5 points out of the 9 total points were allocated to level 1 "application" tasks involving an integral defined function and the inverse function of one of the given functions.

Textbooks included more solicitations for Level 0 or "Recall" tasks (10.7% to 3.1%) while the AP Calculus test included more requests (1.8%) for Level 3 or "Synthesis" tasks (0.1%). This study's results showed that the majority of tasks from both the sampled textbooks and the AP tests fell within the "application" or Level 1 regime of the cognitive complexity model, with the textbooks offering more coverage of application tasks (68.3% versus 62.8%) than the AP Calculus Test. This level was characterized by tasks requiring application of rote procedures typically explicated in the pages immediately preceding the tasks in the textbooks.

Table 7 Textbook Representational Summary in Alpha Order

Source	Word Probs	Algebraic Anchor	Graphical Anchor	Numeric Anchor	Double Anchor	Verbal Target
Anton	86	64 (74.4%)	5 (5.9%)	1 (1.2%)	2 (2.3%)	15 (17.4%)
Finney	39	25 (64.1%)	0 (0.0%)	7 (18.0%)	0 (0.0%)	2 (5.1%)
Forester	52	37 (71.2%)	26 (50.0%)	1 (1.9%)	22 (42.3%)	23 (17.9%)
Hughes-Hallet	85	49 (57.7%)	30 (35.3%)	6 (7.1%)	0 (0.0%)	34 (40.0%)
Larson	114	66 (59.7%)	39 (34.2%)	7 (6.1%)	4 (3.5%)	18 (15.8%)
Rogawski	99	54 (54.5%)	29 (29.3%)	2 (2.0%)	1 (1.0%)	6 (6.0%)
Stewart	51	31 (60.8%)	8 (15.7%)	1 (2.0%)	0 (0.0%)	5 (9.8%)
Thomas	28	20 (71.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	5 (19.5%)
Textbook Totals	555	346 (62.5%)	137 (24.7%)	25 (4.5%)	29 (5.3%)	108 (19.5%)
AP Calculus AB	226	147 (65.0%)	89 (39.4%)	47 (20.8%)	56 (24.8%)	66 (29.2%)

With both the textbooks and the AP tests, the next most popular cognitive complexity was the Interpretation level. This level is characterized by tasks in which a student is required to adapt direct knowledge of calculus concepts into some unfamiliar context. That is, students have to decide which cognitive tool to select to attack a given problem. The uncertainty in the selection of the proper tool is the defining characteristic of the Interpretation level.

Table 8 Textbook Cognitive Complexity Summary in Alpha Order

Source	Word Problems	Level 0 Recall	Level 1 Apply	Level 2 Interpret	Level 3 Synthesize
Anton	86	0 (0.0%)	68 (79.1%)	18 (20.9%)	0 (0.0%)
Finney	39	0 (0.0%)	21 (53.9%)	18 (46.2%)	0 (0.0%)
Forester	52	5 (9.6%)	21 (40.4%)	18 (34.6%)	3 (5.8%)
Hughes-Hallet	85	0 (0.0%)	63 (74.1%)	22 (25.9%)	0 (0.0%)
Larson	114	2 (1.8%)	79 (69.3%)	33 (29.0%)	0 (0.0%)
Rogawski	99	3 (0.03%)	79 (80.0%)	9 (9.1%)	0 (0.0%)
Stewart	51	4 (8.0%)	31 (61.0%)	12 (23.5%)	0 (0.0%)
Thomas	28	3 (10.7%)	17 (60.7%)	6 (21.4%)	0 (0.0%)
Textbook Totals	555	17 (3.1%)	379 (68.3%)	136 (24.5%)	3 (0.01%)
AP Calculus AB	226	0 (0.0%)	142 (62.8%)	83 (36.7%)	4 (1.8%)

There was, roughly, a twelve-percentage point differential between the Level 2 tasks in the textbooks and on the AP test (24.5% versus 36.7%). Phrases like "explain your reasoning" and "justify your answer" were far more common on the AP test than they were in textbooks. This corresponds to the roughly 10% dichotomy between the percentage of solicitations for verbal anchors between textbooks and the tests (19.5% versus 29.2%).

The cognitive complexities of calculus textbook tasks do not align with that of the AP Test. Whereas 24.5% of textbook tasks are of higher order (Level 2 or Level 3), 38.5% of AP tasks were classified as higher order. This is a particularly troubling result with respect to the context of the textbook tasks. Because textbook tasks are located following a section discussing certain concepts, there is a huge implicit hint that the solution to the tasks must somehow invoke the recently covered material. In effect the cognitive complexities of textbook tasks are attenuated by the organization of the textbook. This is not the case on the AP Test. The dichotomy between the complexities of the test and the textbook is actually greater than these numbers indicate.

Qualitative Legitimization

Free Response Item 2007AB3 remains the lowest scoring (0.96 out of 9 possible points) item since 2002 when statistics (College Board, 2012) were available. This item (see Appendix A-2) was framed about a numerical table and included 4 points (College Board, 2012) assigned to analysis and conclusion following an "explain your reasoning " prompt. Specifically these two prompts required the invocation of the Intermediate Value Theorem and Mean Value Theorem, respectively.

Free Response Item 2008AB2 was remarkably similar to #3 from 2007. It too used a numerical table as a stem and it required students to provide analysis and conclusion in response to a "why" question. Unlike the problem from 2007, however, the 2008 problem only allocated 2 points to this type of Level 2 item. The mean score from 2008 was 3.36. For two very similar problems in setup there was a difference in mean score of 2.4 points.

Two points of the differential between the scores between 2007 and 2008 is directly attributable to a lower cognitive complexity required. On the AP test, the difference between the lowest score ever and a mediocre but respectable score is the lessening of cognitive complexity. Perhaps if textbooks would raise the cognitive complexity of its items, the score change could be found in a positive direction.

Explicit Cognitive Attenuation

Based upon 2007 and 2008 results, it is perhaps unsurprising that AP Test creators would shy away from Mean Value Theorem problems. There was no hint of the MVT on the free response questions from either 2009 or 2010, though in 2011 one does appear. Like 2008, the writers relegate the concept to only 2 points (not 4) but this time they go a step further to make the item more accessible. One portion of the 2011 AP Calculus AB #4 problem reads,

> D. Find the average rate of change of f on the interval $-4 \le x \le 3$. There is no point c, $-4 \le c \le 3$, for which $f'(c)$ is equal to the average rate of change. Explain why this statement does not contradict the Mean Value Theorem. (College Board, 2012)

The “why” question is now couched explicitly in terms of the MVT. The author could have asked a different question without giving away the fact that the item was designed for the student to invoke MVT. Instead, there can be no question what cognitive tool a student is supposed to use. The test creators are no longer asking for a Level 2 “interpretation” of the task. Instead they are asking for a Level 0 “recall” of what the MVT is, followed by a Level 1 “application” of the theorem to the given context. Despite the extra hint, the national mean (2.44 out of 9) was actually lower than 2008.

The use of phraseology to lower cognitive complexity is a common technique applied by mathematical problem creators to decrease the “difficulty” for students and increase scores. Though this technique is not as common on the AP test, it is very common in textbooks. In fact, it is the most discernable commonality of the textbook tasks. Sample tasks 1 through 8 (see Appendix B) are taken from each of the sample textbooks. In each task a phrase of the form “Use X to show Y” appears. In each case, the author takes a task from the level 2 “interpretation” realm and relegates it to a level 1 “application” realm.

In effect the author provides a specific cognitive road map for the student to find a solution. There is no need for a student to think about how to solve the problem. The author has given away a solution method. Each task like these is a missed opportunity by the author to empower the student to develop understanding. Why not remove from the textbook all phrases of the form “Use X to show Y” and include them in a teacher ancillary? In that case at least a teacher could decide if a student needed the cognitive assist, rather than giving it to everyone. Many students may not need the hint but all students are given the same cognitive crutch.

Global Conceptualization

In addition to the explicit cognitive attenuation provided by hint or design, as previously noted that, by the very nature of textbook organization into discrete sections focused on specific concepts, there was an implicit cognitive attenuation of textbook tasks. Both the 2007 AP problem and the 2011 AP problem invoke the intermediate value theorem in parallel to the mean value theorem. This parallelism was natural because the theorems are virtually identical conceptually; the former distinguishes a property for a continuous function while the latter considers the same property of the continuous function's derivative.

Despite this natural connection, none of the sample textbooks invoke the intermediate value theorem in the mean value theorem section. None of the textbooks ask the students to distinguish between the two very similar theorems in any of the presented tasks. The students' thinking is guided by the pattern of implicit cognitive attenuation. They do not consider concepts outside of the section's cookbook. Is it any surprise that students struggle when asked to apply global understanding to new situations? This study found no evidence that textbooks deviated from the worn path. Perhaps this explains why a score of 3.36 out of 9 is considered a mediocre score rather than a poor one in the context of the AP calculus test?

Juxtaposition of Representations

Explaining the lower score from the 2011 free response item compared with the 2008 item is challenging in part because the two problems differ in anchor representation. The 2011 item is a graphical representation anchored problem while a table anchors the 2008 item. What is common among the 2008 and 2011 AP items (and also the 2007 item) is a juxtaposition of representational anchors. In each of the three items, students are asked to answer questions through interpretation of multiple representations. All three provide algebraic representations in some aspect in addition to the frame representation, graphical (2011) or numeric (2007 and 2008).

This mixing and matching of representations is the essence of mathematical versatility; it is exactly what the College Board has in mind when it advocates for a rule of four. Students should be able to effortlessly interpret between and among diverse representations of the same scenario. By its design each year's AP Calculus Assessment includes only 6 free response questions, typically each with 3-4 prompts framed by some common representational anchor. Because of brevity of the opportunity to assess students, it was imperative that test designers assess collaborative (rather than discrete) representational understanding.

Calculus textbooks do not have the same issues with brevity. This research found 555 word problems in 15 sections of our 8 sampled textbooks. This is an average of 37 prompts per lesson. These prompts do not include any prompts without a verbal anchor. That is, this number does not include any of the rote exercises that traditionally characterize mathematics homework. The Larson text alone offered an additional 80 rote exercises. As one might expect this research also found an extreme lack of compound representational anchors among textbook tasks. In total, less than 6% of all textbook tasks were set around a compound representational anchor while almost 25% of AP prompts were. In fact, the textbook number is misleading; almost 76% of all doubly anchored textbook tasks came from the Foerster book. Excluding the Forester book, 1.4% of the tasks from the remaining 7 books were doubly anchored. Undoubtedly tasks involving multiple representational anchors are difficult to develop for textbook authors and challenging to solve for students, but in light of the AP assessment it is imperative that such compound representational tasks be included in the textbooks.

Number 33 from page 291 of the Foerster text can illuminate the power of such problems. From the outset the author tells the student that this group of tasks, which begins with a graphical representation, is a proof of Rolle 's Theorem, a corollary to MVT. The first prompt asks students to graph an algebraic function on his calculator. Though this task is low level, it is setting the stage for further robust questions and in the first prompt encompasses an algebraic representation to parallel the given graphical one.

The second prompt asks the student to find a value, a level 1 "application" but this is a different sort of application. Because the author has worked in a second anchor representation, the student now has a choice. He can either find the value algebraically or graphically. This is the choice a developing calculus student needs. The majority of students likely will choose his way, find an answer, and move on. But in this setup the author presents an opportunity for an inquisitive student to connect the two representations. Given the setup of this problem, a teacher can take the opportunity with his class or with a struggling student to tangentially explore the connections between the two representations. Compare this to the situation in which the author's choices preclude deep thought.

The third prompt asks the student to plot a second function on the same screen. It asks the student to compare the current graph to the previous. It requires students to create a table of values for the new function. Here we can demonstrate the power of multiple representations. The author has introduced two more representations, a second graph and a corresponding table. It follows it up by asking the student what do you notice? This is a seemingly innocuously simple question, but it is incredibly profound. The student now has lots of choices. He can verbalize connections he saw between the original graph on the paper and the graph he first made. He can verbalize connections between the second graph he made and the table. He can verbalize connections between the second graph and the first graph. The question is open ended. This question is cognitively amplifying (not attenuating) learning but opening up avenues of thought and discussion.

After developing 4 different representations with the students in previous parts of the problem, the fourth prompt of this problem was the compound question:

> Read the Proof of Rolle's Theorem, which appears in this section. Explain how the work, you've done in this problem relates to the proof. Tell which hypothesis of Rolle's theorem has been mentioned so far in the problem. Is this hypothesis true for the function f? Can the

conclusion of Rolle's theorem be true for a function if the hypotheses aren't? Explain. (Foerster, 2005, p. 219)

For good measure, Forester throws another group of representations into the mix, the verbal description of the proof of Rolle 's Theorem and its accompanying set of graphs and algebraic expressions. None of the nine questions asked across 4 prompts are simple. None give a hint as to how to answer them. Rather than a rudimentary application requiring the most basic understanding of Rolle 's Theorem, the author has chosen to delve into the gory details of the proof encompassing many global concepts not explicitly covered in this section.

It may not be probable that all students will choose to dive into the previously described problem, but at least such an exploration in the wonders of calculus is possible. Most tasks considered by this study do not even offer that opportunity because they lack the representations that make it possible. The ability to create and juxtapose representations, the ability to discuss connections among those representations, and the ability to interpret what those representations mean is the goal of the rule of four.

No Representations

Though lack of compound representations is an issue that plagues textbooks, the issue of lacking of any base representation is also pertinent. The following task in #59 in section 3.2 of the Larson text:

> Two bicyclists begin a race at 8:00 a.m. They both finish the race 2 hours and 15 minutes later. Prove that at some time during the race, the bicyclists are traveling at the same velocity. (Larson et al, 2006, pg, 177)

In its conception this task is not totally dislike some of the tasks we found on the AP Calculus test. It asks students to interpret a mathematical scenario and apply the MVT. In this case, however, there are no anchor representations. There is no graph like in 2011. There is no table like in 2007 or 2008.

This textbook task is of little value to a student because it is too rudimentary and tests no mathematical understanding because by it requires no representational transfer. In effect, the lack of representation is another manner in which an author can lessen the cognitive complexity of a mathematical task. If there is no representation to "interpret" the best a student can do is blindly apply a theorem to a scenario. It is another situation in which the textbook author's choice of presentation has undercut students' ability to learn.

Follow-Up Questions

Questions of the stand-alone variety dominate Calculus Textbooks. Task #59 from Larson lacked any follow-up question; Task #33 from Forester includes a bevy of follow-up questions. These two represent the extremes found in the textbooks. The former is much more common than the later. On page 204, Finney offers a numerically anchored multiple prompt free-response item in #58 (see Appendix B). Due to the succinctness of the presentation, Finney's problem is probably more representative of an AP problem than Foerster, though the learning potential of this one is not quite as rich as the Foerster one. Part A is a basic level 1 "application" and Part B is a straight forward level 2 "interpretation." To this point the problem has stayed with its original numerical table as the lone representation.

Before continuing with the critique of this Finney problem, it should be noted that this is an order of magnitude superior to most of the textbook tasks considered in this study. However, faults were evident in parts C and D. Part C asks the students to find a cubic regression equation for the data given in the table. Cubic Regression is not in the AP Calculus curriculum and this aspect of the problem would likely discourage any teacher from choosing this number for homework. Moreover in Part D, Finney asks the student to use the model to find a formula for the derivative of the original function. There is no scaffolding here. There is no setup question. There is no follow-up question.

A better tact might be for the author to provide the cubic expression. In this fashion, there is no need to pull in an extra-curricular concept and teachers would be more likely to choose this item. Moreover, by introducing the algebraic representation the author can ask a follow-up question eliciting a verbal response and forcing student to address the connections between representations by explicitly asking how they connect.

Alternatively the author could have asked on which interval the function reached its maximum or minimum. This is hinted at in Part B. If the author asked that question, he could follow it up with the current Part B. Then he could follow it up to talk about the rate of change of position and ask when it reaches its maximum and ask for an interpretation. Now in another part the student could attempt to distinguish between the Mean Value Theorem and the Intermediate Value Theorem. Finney missed this opportunity to help his students.

Though this problem has an AP-like setup, it still falls just short of realizing its potential to guide deep understanding. There should be no problem in an AP calculus text that lacks at least one follow-up question. Nine of the 17 textbook prompts considered in this paper lack any sort of follow-up. Follow up questions are a sacrifice calculus textbooks cannot afford to make because the next question is often the more important one.

Task Density

It is hard to imagine completing, in any relatively short period of time (an hour or less), more than a few of the problems like Finney #58 or Foerster #33. Similarly, the AP test does not expect students to plow through problems quickly. The AP test offers a total of 6 stems with approximately 20 prompts each year. These twenty prompts are expected to take a total of 90 minutes or 4.5 minutes per prompt. Assuming each student does 60 minutes of homework a night, consisting of tasks as cognitive complex as the AP test prompts, and takes 4 minutes per task, then a student would be expected to have no more than 15 prompts per night. This equates to approximately 3 AP test stems.

There is a finite amount of time each student will spend on mathematics homework each day. The central question we, as educators face is whether we wish for our students to spend a little time on an each of a large number of problems or a good amount of time on a smaller number of problems. One truism of our classrooms is that quantity and quality of assigned mathematical tasks are inversely proportional.

The implication is that the textbook tasks as presented encourage teachers to assign more easy problems rather than a few challenging problems. This study provides a heuristic to assess the most appropriate tasks from a textbook section. Unfortunately, the creation of “higher quality” homework assignments is not likely possible without an intense and laborious effort by teachers. A set of 15-20 appropriately complex and representationally diverse tasks would be preferable to any set of tasks in a current textbook.

Summary

Textbook tasks should be less dense, avoid cognitive attenuation, move away from the stand-alone item, juxtapose anchor representations, scaffold student solutions, include previously considered overarching concepts and include more profound follow-up questions. In short the cognitive complexity of textbook tasks must be higher. The majority of textbook tasks simply do not go far enough to promote thinking and learning. The tasks are plagued by the aforementioned factors that distract from the opportunity to learn. It is not that students do not learn. Just as has been found before with Algebra I and middle school textbooks (Kulm, 1999), this study finds that students are unlikely to find the opportunity to learn within these textbooks.

Discussion

Is it any surprise that the MVT is at the center of the most bewildering AP Calculus AB free response score ever? The mean value theorem is inextricably a part of calculus' cognitive fabric. It is a formalization of the properties of the derivative over an interval and is the concept atop which the Fundamental Theorem of Calculus is proved. As critical as it is, the Mean Value Theorem is also esoteric, abstract, and not easily accessible to many students nor teachers.

Because the mean value theorem is not as intuitively obvious as other more "common sense" concepts like slope, student learning of it must be properly scaffolded. That is, the right questions must be asked in the right context with the right structure and the right follow-up. Knowledge cannot be transferred from teacher to learner; learning is facilitated by the disequilibrium of challenge not the stasis of rote practice. If we ask the right questions in the right ways students will be empowered to learn in a deep rich conceptual manner, an authentic manner. It is learning that will allow them to construct understanding now and apply it later in a new context.

This study found that textbooks, in general, do not ask the right questions. The American textbook system is stuck in a vicious cycle of good enough. Because established teachers are familiar with traditional textbooks rife with **sterile** tasks, there is no incentive for profit-seeking publishers to make a higher quality textbook despite the thirty-year knell of America's parting mathematical prowess. Administrators, parents, teachers, and students are all comfortable with the sub-standard mathematics education that they know.

For many, if not most educators, the choice of classroom textbook is a proxy for the choice of curriculum. Because the tasks studied here fail to reflect either the AP's prescribed diversity of representations or the cognitive complexity necessary for robust conceptual understanding, nationwide there is little opportunity for authentic learning. Without the improvement of textbooks, our system of mathematics education will not improve.

The discrete nature of the lessons and tasks in American mathematics textbooks must be called into question. What if, instead of a large number of tasks focusing on a small set of concepts, students were presented with a smaller number of tasks encompassing a broader group of concepts? If the problems were less predictable, resilient students would be forced to become more industrious and would be empowered to learn more consequentially.

The one-size-fits-all mentality of publishers has led to mathematics textbooks becoming a series of disconnected lessons. It would be fiscally irresponsible for textbook publishers to include the inter-lesson reflective questions necessary in high-quality textbooks. In order to be palatable to a huge number of school districts and teachers, textbooks cannot impose themselves on reticent students, teachers, and administrators. The path of least resistance (and least quality learning) is the path to the greatest sales. Walter Rudin and other mathematicians knew that it is impossible to mass-market good learning. Quality learning is a function of great teachers, receptive students, and challenging questions with no simple answers.

Teachers, educators, and administrators have, for decades, struggled over the question of how to improve calculus learning? The answer to that question is strikingly simple: improve textbooks. There are thousands of teachers and thousands of schools. There are very few textbooks and even fewer publishers. Why shouldn't textbooks be held to a higher standard? College Board has implemented an Audit system for its teachers and schools. No person or group monitors textbook content. It is time to audit the textbooks. For years textbook publishers have made millions but yet are accountable to no one.

This research began with a question as to why one calculus problem anchored by a table of values caused so much consternation. This research was inspired by a statistic with no precedent and no explanation. It began with a referendum on calculus textbooks and ends with a wholesale indictment. Rudin's *Principles of Mathematical Analysis* is everything that popular calculus textbooks are not. It is short and cogent but more importantly it is efficacious. Modern calculus textbooks are, in fact, just mediocre.

CHAPTER V

SUMMARY AND CONCLUSIONS

New ideas are the lifeblood of education. In the classroom, new ideas are the driving force behind learning. Cognitive dissonance and the struggle to understand drive the assimilation of new concepts in unfamiliar contexts. Research by Dewey, Piaget, Vygotsky, Van Hiele, Brunner, Glasersfield, Shoenfeld, and Cobb has evidenced the importance of assimilating, scaffolding, constructing, and socializing learning. These thinkers have empowered educators to transcend old definitions of learning and understanding. These ideas have encouraged us to embrace a new conceptualization of learning as an active process. No longer must learning be individual, sterile, or furtive. Learning can be open, fertile, and collaborative.

I am currently in my eighth year as a high school teacher and in those eight years I have taught at both public and private schools. I have taught at both urban and suburban schools. I have taught at both new schools and established schools. In those eight years, I have found one common characteristic among my schools and administrators—an absolute aversion to change or challenge of any kind. New ideas are an anathema for educational administrators.

Five years ago, I was in my first year at a suburban public independent school district (ISD). This ISD had spent thousands of dollars for summer professional development provided by the mathematics department from the Exeter Academy. Exeter did not teach mathematics out of textbooks. But rather, their classes were guided by series of single page problem sets intended for small groups. Each problem set was related to the previous set and subsequent set. Each problem set guided systematic simultaneous explorations and discoveries within multiple conceptual threads. The sets were conceptually rich but not overly long. The sets were high quality and low quantity.

Though I did not have the opportunity to attend that summer professional development, I learned of the problem sets later when my veteran department head, a mathematics teacher of some 30 years experience, announced that she had placed binders with the problems in the mathematics work room. The final word she ever said publically about those materials was “feel free to use them in your classes.” I looked at them and thought they were great. My gut inclination was that they were better than textbook tasks I had encountered. I wanted to use them.

About 6 weeks later, I was called to speak to my department head. There had been student complaints about “my” problems, which were actually Exeter problems. I pleaded my case and suggested that the students needed some time to acclimate to the problems. Her response was simple and decisive, “our students are not smart enough for those problems.” I protested vehemently and our conversation ended with her screaming, “Do you want to go talk to the principal?” at me. I responded, “Let’s Go!”

The subsequent conversation with the principal did not go well for me. My stand on curricular principle ended discouragingly. The principal called me childish and arrogant for arguing with a teaching veteran of 30 years. I was not hired to analyze mathematics classes. I was not hired to lead; I was hired to follow. The district had spent a lot of money on textbooks my principal informed me. The problems in our book had been vetted for years. Students from our district in previous years had been “successful” with book problems. Why change now? The mathematical tasks in our textbook were "good enough."

Eighteen months before my principal chastised me for suggesting that textbooks were inadequate, I wondered why my students had fared so poorly with free response item #3 on the 2007 AP Calculus AB examination. That item began with a numerical table and asked students in-part to explain why a certain fact must be true because of the mean value theorem (MVT). Previous research (e.g., Roseman, Kulm, & Shuttleworth, 2001) had suggested that my textbook (Stewart, 2005) might be the culprit. In this study, that suspicion was confirmed.

In the first article, the free response tasks (*n*=60) of all AP Calculus AB-test free-response items available (2002 to 2011) were studied in the context of representational anchors and targets. This article found that the AP's multi-dimensional representational vision, expounded as the rule of four, has not been realized. The article found that students performed less well with numeric anchors and verbal targets. Alternatively graphical targets and anchors produce the best student results. The implication is that the AP's intended curriculum does not align with the curriculum that is implemented by teacher via textbooks in American classrooms.

Given that students do worse with numerical tasks, the logical inference would be that numerical representations are not taught in the classroom and what is taught in the classroom is highly dependent upon the textbook (Kulm & Capraro, 2008). With this background, article 2 sought to create a textbook task assessment instrument. The instrument was based upon previous models developed by van Hiele (1986) and Webb

(1997; 1999) and defined a cognitive complexity scale for calculus textbook tasks in terms of representational transfer. Examples from textbooks and research are offered to define 5 levels: recollection, application, interpretation, synthesis, and abstraction and 6 principles: independence, parsimony, relativism, inheritance, completeness, and reducibility.

In the third article, the previously developed instrument was applied to both AP Calculus AB tasks from 2002 through 2011 and to a sample of textbook tasks from popular calculus textbooks. The textbooks were chosen based upon both the AP Calculus AB suggested textbook list and amazon.com sales rank data. Fifteen sections of the selected textbooks, each centering on the MVT were selected for analysis. All word problems from those sections were evaluated for cognitive complexity and representations invoked. The study found that calculus textbooks underrepresented the numerical anchor and verbal target. It found that the textbooks were both explicitly and implicitly less cognitively complex than the AP test. The article suggested that textbook tasks should be less dense, avoid cognitive attenuation, move away from the stand-alone item, juxtapose anchor representations, scaffold student solutions, incorporate previously considered overarching concepts and include more profound follow-up questions.

This study confirmed the results of Keller and Hirsch (1998) who found that that the anchor representation is associated with student performance; students did poorly with numerical representations because anchors of that type were conspicuously absent from textbook tasks. This analysis found that students perform less well on tasks with numeric anchors and suggests that this pattern of ill performance extends to verbal targets. Alternatively, graphical targets and anchors yielded the best student results. The calculus nation has not achieved multi-representational proficiency and calculus textbooks are unlikely engender deep student learning. This research confirms, just previous research (e.g., Ginsburg, Leinward, Anstrom, & Pollock, 2005) did with other books, that calculus textbooks fail to go beyond exposing students to mathematical mechanics of routine problems.

The story of calculus pedagogy is the story of a culture war between the antiquated and modern pedagogies. The mathematics wars of years past (Schoenfeld, 2004) have not disappeared; a cold pedagogical war is fought everyday in classrooms around the country (Kaplan & Kaplan, 2008; Lockhart, 2009). The pedagogy espoused in classrooms is decidedly not cognitively guided. The evidence for this fact, which was previously discussed anecdotally (St. Jarre, 2008), is now more formally documented in terms of a representation-based cognitive-complexity analysis of major textbook tasks.

As I learned in one suburban high school, antiquated orthodoxy is entrenched, reinforced, and buttressed. I walked into a high school hoping to teach great ideas through great problems. An ambitious principal whose future was dependent upon maintaining the status quo, an entrenched veteran teacher whose primary concern was minimizing personal effort, and a publisher whose number one priority was profits -- dashed my hopes. All rationale used against the allowance of more challenging problems in one classroom was based in the unchallenged perception of the infallibility of the textbook.

How many textbook adoption processes across the nation are rigorous and based on high quality research? The process of textbook analysis and adoption must cease to be intellectually soft (Battista & Clemens, 2000). The choice of textbook cannot just be a show of hands. As a colleague of mine once remarked, the choice of Larson calculus in his school district was a choice to minimize prep work. They had always used the Larson text, changing would only cause more work for he and his fellow teachers. What was the point? The point is that the question of quality student learning never factored into the process.

The 2006-2007 school year ended for me with a conundrum about an AP Calculus AB test item. I now know why my students (and the rest of their counterparts across the nation) faired so poorly on that question. I never asked the hard questions; the hard questions about the validity of my textbook, which meant I did not have the hard questions available to ask of my students, who in turn had no hard questions to ask of me. The cycle of learning was stunted because I trusted my textbook. I was wrong. It is time to analyze textbooks, all of them. If the College Board can develop a formal review process for its teachers' syllabi, why can't it institute a review of the dozen textbooks it recommends? Why is that teachers are held to a higher standard than textbooks? There are thousands of teachers. There are very few textbooks.

Just as students must be challenged intellectually with non-routine thought provoking problems, our teachers and curriculum selectors must seek a thought-provoking process of textbook analysis. It is time we require more of the publishers who make millions proliferating pedagogical garbage. We live in a mathematics culture where textbooks that lack hard questions are hardly questioned by teachers, who implicitly trust the publishers, about whom the hard questions are never asked. All calculus textbooks are mediocre because we have never demanded better.

In the past 6 six years, Texas A&M University has not graduated a single engineering Ph.D. who did not take calculus on the university level (Nite, 2012). In other words, no student who has taken AP Calculus at the high school level has earned an engineering doctorate degree. For a university, that as of Fall 2011, boasted more than 2800 graduate level engineering students, that fact is an astonishing indictment of the AP curriculum. The AP curriculum is not satisfactorily empowering students to attain quality calculus understanding.

More than 2 million students have taken the calculus AB examination over the past decade. In that same time, more students have sat for the BC examination and still more have taken the class but opted not to take either examination. The number of students who have taken the exams each year has been steadily rising. The College Board is successfully marketing and expanding its curricular reach. But are those gains being attained at the expense of quality learning? The findings of this dissertation and the previous research cited certainly do not controvert that hypothesis.

In recent years, the College Board has developed an auditing process for its calculus teachers. Across the nation thousands of calculus teachers are required to submit their syllabi for review. The syllabi are checked to assure that classes meet AP quality standards. Meanwhile, the AP publishes, though does not endorse, a list of calculus textbooks on its website (College Board, 2012). These books are written by a select few authors and disseminated by even fewer publishers. Whereas the College Board has made an effort to crack down on rogue teachers, it has made no such effort with textbooks. Publishers make millions from the AP course with impunity.

The College Board needs to vet its calculus textbooks. It needs to develop a process through which books can be assessed for both strengths and weaknesses. Ultimately the only way in which to change textbooks is to change the market for those textbooks. The College Board has put out a number of supplemental resources in the past on curricular issues, including one on reasoning with tabular data in the past. Materials on topics such as reasoning with multiple representations or choosing good mathematical tasks would be monumentally helpful to both teachers and textbook authors.

Adjustments are necessary. Students are slipping through the cracks and we are losing the STEM graduates that we need to assure our culture's future prosperity. Every aspect of the system of calculus education could be better, particularly the textbooks on atop which the AP curriculum is implemented. For the sake of the students who will learn in the STEM-dominated 21^{st} century, the textbooks cannot be just mediocre.

REFERENCES

Allendoerfer, C. (1963). The case against calculus. *Mathematics Teacher, 56*. 482-485.

Anton, H., Bivens, I., & Davis, S. (2005). *Calculus: Early transcendentals single variable*. (8th ed.). Hoboken, NJ: Wiley.

Arbaugh, F., & Brown, C. A. (2004). Worthwhile mathematical tasks: Designing a learning tool for high school mathematics teachers. In R. Rubenstein (Ed.), *Effective mathematics teaching: Implementing the teaching principle*. (pp. 27-41). Reston, VA: NCTM.

Aspinwall, L., & Shaw, K. L. (2002). Representations in calculus: Two contrasting cases. *Mathematics Teacher*, *95*, 434-39.

Baker, B., Cooley, L., & Trigueros, M. (2000). A calculus graphing schema. *Journal for Research in Mathematics Education, 31, 557-578.*

Battista, M. T., & Clements, D. H. (2000). Mathematics curriculum development as a scientific endeavor. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 737-760). Mahwah, NJ: Erlbaum.

Battista, M. T. (1999). The mathematical miseducation of America's youth: Ignoring research and scientific study in education. *Phi Delta Kappan, 80,* 424-33

Black, P., Harrison, C., Lee, C., Marshall, B., &William, D. (2004). Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan, 86*, 8-21.

Braiden, D. (2011). To solve or not to solve, that is the problem. *Australian Senior Mathematics Journal*, *25*(2), 7-13.

Bressoud, D. (2009). *AP® calculus: What we know*. Retrieved from Mathematical Association of America Web site: http://www.maa.org/columns/launchings/launchings_06_09.html

Bruner, J. (1960). On learning mathematics: What places do discovery, intuition and translation have in the learning of mathematics? Are our present activity and efforts to improve mathematics teaching the beginning of an educational renaissance? *Mathematics Teacher*, *53*. 610-619.

Cai, J., & Lester, F. K. (2005) Solution representations and pedagogical representations in Chinese and U.S. classrooms. *Journal of Mathematical Behavior, 24*. 221-237.

Caldwell, J. H., & Goldin, G. A. (1987). Variables affecting word problem difficulty in secondary school mathematics. *Journal for Research in Mathematics Education*, *18*(3), 187-196.

Capraro, M. M., Yetkiner, E., Ozel, S. Capraro, R., Ye, S., & Kim, H. (2009, April). *An international perspective on sixth graders' interpretation of the equal sign.* Presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Capraro, M. M. (2001, February). *Defining constructivism: Its influence on the problem solving skills of students*. Paper presented at the Annual Meeting of the Southwest Educational Research Association, New Orleans, LA.

Capraro, R. M., Capraro, M. M., Younes, R., Han, S. Y., & Garner, K. (2012). Changes in equality problem types across four decades in four second and sixth grade textbook series. *Journal of Mathematics Education*, *5*(1), 166-189.

Capraro, R. M., & Slough, S. W. (2008). *Project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach*. Rotterdam, Netherlands: Sense.

Capraro, R. M., & Yetkiner, Z. E. (2008). Teachers' role in developing representational fluency in middle grades. In G. Kulm (Ed.) *Teacher knowledge and practice in middle grades mathematics*. (pp. 273-286). Rotterdam, Netherlands: Sense.

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann.

Cheng, P. (1999). *Representational analysis and design: What makes an effective representation for learning probability theory?* Technical Report No. 65. ESRC Centre for Research in Development, Instruction, and Training. Department of Psychology, University of Nottingham, UK.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education*, *23*(1), 2-33.

College Board. (2012). *AP calculus AB course home page*. Retrieved from http://apcentral.collegeboard.com/apc/public/courses/teachers_corner/2178.html

Crawford, D. B., & Snider, V. E. (2000). Effective mathematics instruction and the importance of curriculum. *Education and Treatment of Children*, *23*, 122-142.

Cunningham, R. F. (2005). Algebra teachers' utilization of problems requiring transfer between algebraic, numeric, and graphic representations. *School Science and Mathematics, 105*(2), 73-79.

Curcio, F.R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382-393.

Darling-Hammond, L. (1998). Unequal opportunity: Race and Education. *Brookings Review*, *16*,(2), 28-32.

Demana. F. (1994). College precalculus courses: Challenges and changes. In A. Solow (Ed.). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America.

Douglas, R.G. (Ed.). (1986). *Toward a lean and lively calculus.* Washington, D.C: Mathematical Association of America.

Dreyfus, T., & Eisenberg, T. (1982). Intuitive functional concepts: A baseline study on intuitions. *Journal for Research in Mathematical Education*, *6*(2), 18-24.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking, in D. Tall (Ed.), *Advanced Mathematical Thinking* (pp.95-126), Dordrecht, Netherlands: Kluwer.

Edens, K., & Potter, E. (2008). How students "unpack" the structure of a word problem: Graphic representations and problem solving. *School Science and Mathematics, 108*(5), 184-196.

Ferrini-Munday, J., & Graham, K. (1991). Overview of the calculus curriculum reform effort: Issues for learning, teaching, and curriculum Development. *American Mathematical Monthly*, *98*, 627-35

Finney, R. L., Demana, F. D., Waits, B. K., & Kennedy, D. (2007). *Calculus: Graphical, numeric, algebraic* (3rd ed.). Boston, MA: Pearson Prentice Hall.

Foerster, P. A. (2005). *Calculus: Concepts and applications* (2nd ed.). York, PA: Key.

Freudenthal, H. (1973). *Mathematics as an education task.* Dordrecht, Netherlands: Reidel.

Geiser, D. B. (2008). *First ever ap course audit benefits stakeholders*. Retrieved from http://www.tip.duke.edu/node/896

Ginsburg, A., Leinwand, S., Anstrom, T., & Pollock, E. (2005). *What the United States can learn from Singapore's World Class Mathematics System and what Singapore can learn from the United States: An exploratory study.* Washington, D.C. American Institutes for Research.

Goldin, G. A. (2003). Representation in school mathematics: A unifying research perspective. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *Research companion to principles and standards for school mathematics* (pp. 275-284). Reston, VA: National Council of Teachers of Mathematics.

Goldin, G. A., & Shteingold, N. (2001). Systems of representations and the development of mathematical concepts. In A. A. Cuoco & F. R. Curcio (Eds.) *The roles of representation in school mathematics* (pp. 1-23). Reston: VA. National Council of Teachers of Mathematics.

Gordon, F.S. (2006). Assessing what students learn: Reform versus traditional precalculus and follow-up calculus. In N. B. Hastings, F. S. Gordon, S. P. Gordon, & J. Narayan (Eds.), *A fresh start for collegiate mathematics: Rethinking the courses below calculus* (pp. 181-192). Washington, DC: Mathematical Association of America.

Gordon, S., & Hughes Hallett, D. (1994). Lessons from the calculus reform effort for precalculus reform. In A. Solow (Ed.). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America.

Grassl, R. M., & Mingus, T. T. Y. (2002). On the shoulders of technology: Calculators as cognitive amplifiers. *International Journal of Mathematical Education in Science and Technology*, *33*(5), 715-723.

Gregg, J. (1995). The tensions and contradictions of the school mathematics curriculum. *Journal for Research in Mathematics Education*. *26(5)*, 442-466.

Halmos, P. R. (1980). Heart of mathematics. *American Mathematical Monthly, 87*, 519-524.

Hawking, S.W. (2001). *Universe in a nutshell*. New York, NY: Bantam

Heid, M. K. (1988). Resequencing skills and concepts in applied calculus using the computer as a tool. *Journal for Research in Mathematics Education, 19*(1). 3-25.

Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second grade arithmetic. *American Educational Research Journal, 30,* 393-425.

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. 65-97). New York, NY: McMillan.

Hoffer, A. (1983). Van Hiele-based research. In R. Lesh &M. Landau (Eds.), *Acquisition of mathematics concepts and processes*. Orlando, FL: Academic Press.

Hughes Hallett, D. (2006). The importance of a story line: functions as models of change. In N. B. Hastings, F. S. Gordon, S. P. Gordon, & J. Narayan (Eds.), *A fresh start for collegiate mathematics: Rethinking the courses below calculus.* (pp. 352-354). Washington, DC: Mathematical Association of America.

Hung, D. W. L., & Chen, D.-T. (2001). Situated cognition, Vygotskian thought and learning from the communities of practice perspective: Implications for the design of web-based e-learning. *Education Media International, 38*(1), 3-12.

Janvier, C. (1987). Representation and understanding: The notion of function as an example. In C. Janvier (Ed.) *Problems of representation in the teaching and learning of mathematics*, (pp. 67-72). Hillsdale, NJ: Erlbaum.

Johnson R. B. & Onwuegbuzie A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher 33*(7),14–26.

Kajander, A., & Lovric, M. (2009). Mathematics textbooks and their potential role in supporting misconceptions. *International Journal of Mathematical Education in Science and Technology*, *40*(2), 173-181.

Kaplan, R., & Kaplan, E. (2008). *Out of the Labyrinth: Setting mathematics free*. New York: Oxford.

Kaput, J. (1985). Representation and problem solving: Methodological issues related to modeling. In E. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives*, (pp. 381-398). Hillsdale, NJ: Erlbaum.

Kaput, J. (1987). Representation systems and mathematics. In C. Janvier (Ed.) *Problems of representation in the teaching and learning of mathematics*, (pp. 19-26). Hillsdale, NJ: Erlbaum.

Keller, B. A., & Hirsch, C. R. (1998). Student preferences for representations of functions. The *International Journal of* Mathematical *Education* in *Science* and *Technology, 29* (1), 1-17.

Knoebel, A., Kurtz, D. S., & Pengelley, D. (1994). A case study of a partnership of equals: Calculus meets precalculus. In A. Solow (Ed.). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America.

Knuth, E. J. (2000). Student understanding of the Cartesian connection: An exploratory study. *Journal for Research in Mathematics Education, 31*(4), 500-508.

Kulm, G. (1994). New directions for mathematics assessment. In G. Kulm, (Ed.), *Assessing higher order thinking in mathematics*. Hillsdale, NJ: Erlbaum.

Kulm, G. (1999). *Evaluating mathematics textbooks*. Retrieved from http://www.project2061.org/publications/articles/articles/cbe.htm

Kulm, G., & Capraro, R. M. (2008). Textbook use and student learning of number and algebra ideas in middle grades. In G. Kulm (Ed.) *Teacher knowledge and practice in middle grades mathematics* (pp. 255-272). Rotterdam, Netherlands: Sense.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery.* New York, NY: Cambridge University Press.

Lampert, M., & Cobb, P. (2003). Communication and language. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *Research companion to principles and standards for school mathematics* (pp. 275-284). Reston, VA: National Council of Teachers of Mathematics.

Lappan, G. & Briars, D. (1995). How should mathematics be taught? In I. M. Carl (Ed), *Prospects for school mathematics* (pp. 131-156). Reston, VA: National Council of Teachers of Mathematics.

Larson, R., Boswell, L., & Stiff, L. (2001). *Geometry*. Boston, MA: McDougal Littell.

Larson, R., Hostetler, R. P., & Edwards, B. H. (2006). *Calculus* (8th ed.). Boston, MA: Houghton Mifflin.

Lesh, R. (2006). *New directions for research on mathematical problem solving.* Keynote address of the 29th annual Conference of the Mathematics Education Research Group of Australasia, Canberra, Australia.

Lesh, R., & Yoon, C. (2004). Evolving communities of mind – In which development involves several interacting and simultaneous developing strands. *Mathematical Thinking and Learning, 2,* 205-226.

Lesh, R., Hoover, M., Hole, B., Kelly, E., & Post, T. (2000). *Principles for developing thought revealing activities for students and teachers*. Mahwah, NJ: Erlbaum

Lester, F. K. & Kehle, P. E, (2003). From problem solving to modeling: The evolution of thinking about research on complex mathematical activity. In R. A. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 501-518). Mahwah, NJ: Erlbaum.

Lockhart, P. (2009). *Mathematician's lament: How school cheats us out of our most fascinating and imaginative art form*. New York, NY: Bellevue.

Manouchehri, A., & Lapp, D. A. (2003). Unveiling student understanding: The role of questioning in instruction. *Mathematics Teacher*, *96*(8), 562-566.

Manouchehri, A., & St. John, D. (2006). From classroom discussions to group discourse. *Mathematics Teacher*, *99*(8), 544-551.

McCrory, R. (2006). Mathematicians and mathematics textbooks for prospective elementary teachers. *Notices of the AMS, 53*(1), 20-29.

McNeely, M. E. (1997). *Guidebook to examine school curricula: TIMSS as a starting point*. (ERIC Document Reproduction Service No. ED410128)

Muir, T., Beswick, K., & Williamson, J. (2008). "I'm not very good at solving problems": An exploration of students' problem solving behaviors. *Journal of Mathematical Behavior*, *27*, 228-241.

Mullis, I.V.S. (1996). Attaining excellence: TIMSS as a starting point to examine mathematics assessments. An in-depth look at geometry and algebra. Washington, DC: National Center for Education Statistics, Office of Education Research and Improvement, U.S. Dept. of Education

National Council of Teachers of Mathematics. (1980). *An agenda for action: Recommendations for school mathematics of the 1980s*. Reston: VA: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards*. 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. (2009). *Focus on high school mathematics: Reasoning and sense making*. Reston: VA: Author.

Nelissen, J. M. C., & Tomic, W. (1996). *Representation and cognition.* (ERIC Document Reproduction Service No. ED402012).

Nite, S. B. (2012). Bridging secondary mathematics to post-secondary calculus: A summer bridge program. Unpublished dissertation. College Station, TX: Texas A&M.

Pandiscio, E., & Orton, R.E. (1998). Geometry and metacognition: An analysis of Piaget's and Van Hiele's perspectives. *Focus on the Learning Problems in Mathematics*, *20*(2&3), pp. 78-87.

Pegg, J., & Davey, G. (1998). Interpreting student understanding in geometry: A synthesis of two models. In R. Lehrer & D. Chazan (Eds.) *Designing learning environments for developing understanding of geometry and space* (pp. 109-135). Mahwah, NJ: Erlbaum.

Piccolo, D. L., Carter, T. A., Harbaugh, A. P., Capraro, M. M., & Capraro, R. M. (2008). Quality of instruction: Examining discourse in middle school mathematics instruction. *Journal of Advanced Academics, 19*, 376-410.

Polya, G. (1967). *Mathematical discovery: On understanding, learning, and teaching problem solving*. Hoboken, NJ: Wiley.

Polya, G. (1973). *How to solve it. Princeton,* NJ: University Press.

Porzio, D. (1999). Effects of differing emphases in the use of multiple representations and technology in students' understanding of calculus concepts. *Focus on Learning Problems in Mathematics, 21*(3), 1-29.

Reys, R., Reys, B., Tarr, J., & Chavez, O. (2006). Assessing the impact of standards-based middle school mathematics curricula on student achievement and the classroom learning environment. Washington, DC: National Center for Education Research.

Rogawski, J. (2008). *Calculus: Early transcendentals*. New York, NY: W.H. Freeman.

Romberg, T.A., Fennema, E., & Carpenter, T. P. (1993). *Integrating research on the graphical representation of functions.* Mahwah, NJ: Erlbaum.

Romero, C. (2008). Who killed Bob Krusty: A dynamic problem solving event. In R.M. Capraro & S.W. Slough (Eds.) *Project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach* (pp. 182-183). Rotterdam, Netherlands: Sense.

Romero, C. (2012). Pulling back the curtain on secondary calculus: A representational analysis of AP Calculus AB Free Response items, 2002-2011. Unpublished dissertation. College Station, TX: Texas A&M.

Romero, C. (2012). What I cannot create I do not understand: Using representations to diagnose the cognitive complexity of mathematics tasks. Unpublished dissertation. College Station, TX: Texas A&M.

Roseman, J. E., Kulm, G., & Shuttleworth, S. (2001). Putting textbooks to the test. *ENC Focus, 3*, 56-59.

Rudin, W. (1997). *The way I remember it: History of mathematics* (Vol. 12). Providence, RI : American Mathematical Society

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1996). *A splintered vision: An investigation of US science and mathematical education.* Retrieved http://ustimss.msu.edu/splintrd.pdf

Schnepp, M. J., & Nemirovsky, R. (2001). Constructing a foundation for the fundamental theorem of calculus. In A. Cuoco & F. Curcio (Eds.), *The roles of representation in mathematics* (pp. 90-102). Reston, VA: NCTM.

Scher, D., & Goldenberg, E. P. (2001). A multirepresenational journey through the law of cosines. In A. Cuoco & F. Curcio (Eds.), *The roles of representation in school mathematics* (pp. 117-128). Reston, VA: NCTM.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.). *Cognitive science and mathematics education*. (pp. 189-215). Hillsdale, NJ: Erlbaum.

Schoenfeld, A.H. (1988) When good teaching leads to bad results: The disasters of “well-taught” mathematics courses, *Educational Psychologist*, *23*(2), 145-66.

Schoenfeld, A.H. (2004). The math wars. *Educational Policy*, *18*, 253-286.

Schwartz, M (2006). For whom do we write the curriculum?. *Journal of Curriculum Studies*, *38*, 449-457.

Selden, A., Selden, J., Hauk, S., & Mason, A. (1999). *Do calculus students eventually learn to solve non-routine problems* (Technical Report No. 1999-5). Cookeville, TN: Tennessee Technological University.

Selden, A., Selden, J., Hauk, S., & Mason, A. (2000). Why can't calculus students access their knowledge to solve non-routine problems? *CBMS Issues in Mathematics Education, 8,* 128-153.

Shaughnessy, M., & Burger, W. F. (1985). Spadework prior to deduction in geometry. *Mathematics Teacher, 78*, 419-427.

Sieben, L. (2011, Feburary 9). *Number of ap test takers has nearly doubled since 2001*. Retrieved from http://chronicle.com/article/Number-of-AP-Test-Takers-Has/126313/

Smith III, J. P., & Star, J.R. (2007). Expanding the notion of impact of K-12 standards based mathematics and reform calculus programs. *Journal for Research in Mathematics Education, 38*, 3-34.

Smith, D. A. (1994). Trends in calculus reform. In A. Solow (Ed.). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America.

Smith, M. S., & Stein, M. K. S. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3,* 344-350.

Solow, A. (Ed.). (1994). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America

St. Jarre , K. (2008). They knew calculus when they left: The thinking disconnect between high school and university. *Phi Delta Kappan*, *90*, 123-126.

Stacey, K. (2005). The place of problem solving in contemporary mathematics curriculum documents. *Journal of Mathematical Behavior, 24,* 341–350.

Stanley, S. (2002). Revitalizing pre-calculus with problem-based learning. *Journal of General Education*, *51,* 306-315.

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

Stewart, J. (2005). *Calculus: Concepts and contexts* (3rd ed.). Belmont, CA: Thomson Brooks/Cole.

Stigler, S. M. (1986). *The history of statistics: the measurement of uncertainty before 1900*. Cambridge, MA: Belknap.

Styliianou, D. A., & Silver, E. A. (2004). The role of visual representations in advanced mathematical problem solving: An examination of expert-novice similarities and differences. *Mathematical Thinking and Learning*, *6*(4), 353-387.

Tall, D. (1992). Mathematical processes and symbols in the mind. In Z. A. Karian (Ed.), Symbolic computation in undergraduate mathematics education. MAA Notes 24. *Mathematical Association of America, 24,* 57-68.

Tall, D. (1997). Functions and calculus. In A. J. Bishop et al. (Eds.), *International Handbook of Mathematics Education*. (pp. 289–325). Dordrecht, Netherlands: Kluwer.

Thomas, G., Weir, M. D., Hass, J., & Giordano, F. R. (2008). *Calculus: Early transcendentals*. (11th ed.). Boston, MA: Pearson Addison Wesley.

Thornton, S. (1998). Constructing geometric reasoning. *Australian Mathematics Teacher, 54* (3). 6-11.

Tucker, T. W. (Ed.). (1990). *Priming the calculus pump: Innovations and resources*. Washington, DC: Mathematical Association of America.

Tucker, T. W. (1996). Nonalgebraic approaches to calculus. In A.W. Roberts (Ed.) *Calculus: The dynamics of change.* (pp. 16-19). Washington, DC: Mathematical Association of America.

U.S. Department of Education. (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry*. Washington, DC: National Inst. of Education. (ERIC Document Reproduction Service No. ED 220288)

Usiskin, Z. (1994). Recent developments in secondary school mathematics and their implications. In A. Solow (Ed.), *Preparing for a new calculus* (MAA Notes #36, pp. 21-27). Washington, DC: Mathematical Association of America.

Usiskin, Z. (1997). Applications in the secondary school mathematics curriculum: A generation of change. *American Journal of Education*, *106*, 62-84.

Van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education.* Orlando, FL: Academic Press.

Vygotsky, L., & Kozulin, A. (1986). *Thought and language*. Cambridge, MA: MIT.

Waits, B. K., & Demana, F. (1994). Graphing calculator intensive calculus: A first step in calculus reform for all students. In A. Solow (Ed.). *Preparing for a new calculus: Conference proceedings*. Washington, DC: Mathematical Association of America.

Webb, N. L. (1997). *Research Monograph Number 6: Criteria for alignment of expectations and assessments on mathematics and science education*. Washington, DC: CCSSO.

Webb, N. L. (1999). *Research Monograph No. 18: Alignment of science and mathematics standards and assessments in four states*. Washington, DC: CCSSO.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*, 7-25.

Witzel, B. S., & Riccomini, P. J. (2007). Optimizing math curriculum to meet the learning needs of students. *Preventing School Failure*,*52*, 13-18.

Yan, Z., & Lianghuo, F. (2004*). The representation of problems in Chinese and US mathematics textbooks.* Retrieved from www.icme-organisers.dk/dg14/DG14-Zhu%20&%20Fan.pdf

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179-217.

APPENDIX A

SAMPLE AP CALCULUS AB FREE RESPONSE ITEMS

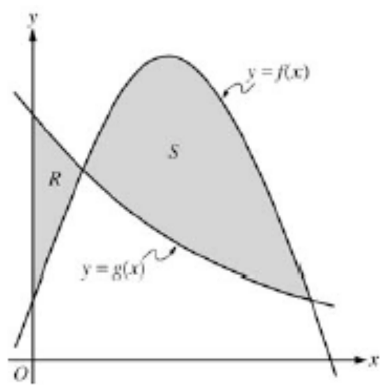
1. AP Free Response Question #1 from 2005 (National Mean:5.73/9)

Let f and g be functions given by $f(x) = \frac{1}{4} + \sin(\pi x)$ and $g(x) = 4^{-x}$. Let R be the shaded region in the first quadrant enclosed by the y-axis and the graphs of f and g, and let S be the shaded region in the first quadrant enclosed by the graphs of f and g as shown in the figure.

A. Find the area of R [3 points].

B. Find the area of S [3 points].

C. Find the volume of the solid generated when S is revolved about the horizontal line $y = -1$.[3 points]



2. AP Free Response Question #3 from 2007 (National Mean: 0.96/9)

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
1	6	4	2	5
2	9	2	3	1
3	10	-4	4	2
4	-1	3	6	7

The functions f and g are differentiable for all real numbers and g is strictly increasing. The table above gives values of the functions and their first derivatives at selected values of x. The function h is given by $h(x) = f(g(x)) - 6$.

A. Explain why there must be a value r for $1 < r < 3$ such that $h(r) = -5$. [2 points]

B. Explain why there must be a value c for $1 < c < 3$ such that $h'(c) = -5$. [2 points]

C. Let w be the function given by $w = \int_1^{g(x)} f(t)dt$. Find the value of $w'(3)$. [2 points]

D. If g^{-1} is the inverse function of g, write an equation for the line tangent to the graph of $y = g^{-1}(x)$ at $x = 2$. [3 points]

3. AP Free Response Question #2 from 2008 (National Mean: 3.36/9)

t (hours)	0	1	3	4	7	8	9
$L(t)$ (people)	120	156	176	126	150	80	0

Concert tickets went on sale at noon ($t = 0$) and were sold out within 9 hours. The number of people waiting in line to purchase tickets at a time t is modeled by a twice-differentiable function L for $0 \le t \le 9$. Values of $L(t)$ at various times t are show in the table above.

A. Use the data in the table to estimate the rate at which the number of people waiting in line was changing at 5:30 p.m. ($t = 5.5$). Show the computations that lead to your answer. Indicate units of measure. [2 points]

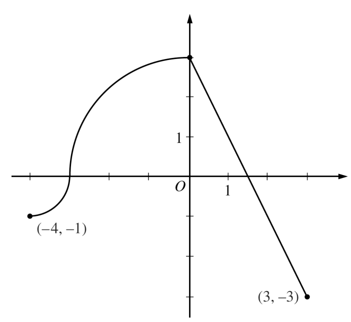
B. Use a trapezoidal sum with three subintervals to estimate the average number of people waiting in line during the first four hours of the sale. [2 points]

C. For $0 \le t \le 9$, what is the fewest number of times at which $L'(t)$ must be equal to 0. Give a reason for your answer. [3 points]

D. The rate at which tickets were sold for $0 \le t \le 9$ is modeled by $r(t) = 550te^{-t/2}$ tickets per hour. Based on the model, how many tickers were sold by 3 p.m. ($t = 3$), to the nearest whole number? [2 points]

4. AP Free Response Question #4 from 2011 (National Mean: 2.44/9)

The continuous function f is defined on the interval $-4 \le x \le 3$. The graph consists of two quarter circles and one line segment as show in the figure. Let $g(x) = 2x + \int_0^x f(t)dt$.



Graph of f

A. Find $g(-3)$. Find $g'(x)$ and evaluate $g'(-3)$. [3 points]
B. Determine the x-coordinate of the point at which g has an absolute maximum on the interval $-4 \le x \le 3$. Justify your answer. [3 points]

C. Find all values of c on the interval $-4 \le x \le 3$ for which the graph of g has a point of inflection. Given a reason for your answer. [1 point]

D. Find the average rate of change of f on the interval $-4 \le x \le 3$. There is no point c, $-4 \le c \le 3$, for which $f'(c)$ is equal to the average rate of change. Explain why this statement does not contradict the Mean Value Theorem. [2 points]

APPENDIX B

SAMPLE CALCULUS TEXTBOOK TASKS

1. Use the inequality $\sin(x) \le x$ which holds for $x \ge 0$ to find an upper bound for the value of $\int_0^1 \sin(x)dx$. (Finney, p. 292, #42)

2. Let $0 < a < b$. Use the Mean Value Theorem to show that $\sqrt{b} - \sqrt{a} < \frac{b-a}{2\sqrt{a}}$ (Larson, p. 178 #85)

3. Show that $\tan(x) > x$ for $(0,\pi/2)$. [Hint: Show that $f(x) = \tan(x) - x$ is increasing on $(0,\pi/2)$] (Stewart, p. 289, #51)

4. Use the Racetrack principle and the fact that $\sin 0 = 0$ to show that $\sin(x) \le x$ for all $x \ge 0$ (Huges-Hallett, p 158, #10)

5. Use the Constant Difference Theorem (5.7.3) to show that if $f'(x) = g'(x)$ for all x on $(-\infty,\infty)$ and if f and g have the same value at some point x_0, then $f(x) = g(x)$ for all x on $(-\infty,\infty)$ (Anton, p 335, #24a)

6. Suppose that $f(0) = 4$ and $f'(x) \le 2$ for $x > 0$. Apply the MVT to the interval $[0,3]$ to prove that $f(3) \le 10$. (Rogawski, p 237, #62)

7. Use the same derivative argument as was done to prove the Product and Power Rules for logarithms, to prove the Quotient Rule property (Thomas, p. 284 #65)

8. Prove that if an object moves with a constant acceleration, such as it does in ideal free fall, then its average velocity over the time interval is the average of the velocities at the beginning and end of the interval. (This result leads to one of the physics formulas you may have learned and that may have led to you to a false conclusion about average velocity when the acceleration is not constant). (Foerster, p 512, #13)

9. The proof of Rolle's theorem shows that a high point $f(c)$ for the open interval (a,b) the difference quotient $\frac{f(x)-f(c)}{x-c}$ is always positive (or zero) when $x<c$ and always negative with $x>c$. In this problem you will show graphically and numerically that this fact is true for a fairly complicated function.

a. The figure below shows the graph of $f(x)=25-(x-5)^2+4\cos[2\pi(x-5)]$. Plot the graph as y_1. Does your graph agree with the figure?
b. Find $f'(x)$. How is the value of $f'(5)$ consistent with the fact that the high point of the graph is at $x=5$?
c. Let y_2 be the difference quotient $y_2=\frac{y_1-f(5)}{x-5}$. Plot y_2 on the same screen as y_1. Sketch the result. Then make a table of values of the difference quotient from each 0.5 unit of x from $x=3$ to $x=7$. What do you notice about the table and the graph about the values of y_2 for $x>5$ and $x<5$?
d. Read the Proof of Rolle's Theorem, which appears in this section. Explain how the work, you've done in this problem relates to the proof. Tell which hypothesis of Rolle's theorem has been mentioned so far in the problem. Is this hypothesis true for the function f? Can the conclusion of Rolle's theorem be true for a function if the hypotheses aren't? Explain.

(Foerster, p 219 #33)

10. Priya's distance D in meters from a motion detector is given in the table below.

a) Estimate when Priya is moving toward the motion detector and away from the motion detector
b) Give an interpretation of the extreme values in terms of this problem situation
c) Find a cubic regression equation D=f(t) for the data and superimpose it on a graph of the scatter plot of the data
d) Use the model in C for f to find a find a formula the derivative of F. Use your formula to estimate the answers to part a.

(Finney, p. 204, #58)

APPENDIX C

COGNITIVE COMPLEXITY MODEL

<table>
<tr><th></th><th>#</th><th>Verb</th><th>Reps</th><th colspan="2">Indicators/Description</th></tr>
<tr><td rowspan="4">Exercises</td><td rowspan="2">0</td><td>Recall</td><td>0</td><td colspan="2">0A. Recall the definition of rectangle
0B. Identify the number of roots of this graph</td></tr>
<tr><td></td><td colspan="3">• No cognitive effort is extended, this level is one of habitual response</td></tr>
<tr><td rowspan="2">1</td><td>Apply</td><td>1</td><td colspan="2">1A. Rote Procedures (Find Solutions, Simplify Expressions, Compute)
1B. Simple Data Operations (Observe, Collect, Display, Compare)</td></tr>
<tr><td></td><td colspan="2">• Selection of tool is not required of students; tasks are often rote procedure</td><td>• A single representation: Algebraic, Numeric, or Graphical</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">2</td><td>Interpret</td><td>2+</td><td colspan="2">2A. Explain your reasoning
2B. Argue logically from evidence
2C. Explain phenomena in terms of concepts</td></tr>
<tr><td></td><td colspan="2">• Planning Required
• More than 1 solution possible</td><td>• Excessive of “grunt” work does not entail higher levels, higher cognitive demand does</td></tr>
<tr><td rowspan="4">Problems</td><td rowspan="2">3</td><td>Synthesize</td><td>3-4</td><td colspan="2">3A. Prove Conjecture
3B. Connect related concepts & phenomena
3C. Synthesize or Generalize ideas into new concepts</td></tr>
<tr><td></td><td colspan="2">• Multiple Possible Paths to Solutions
• Students might dislike lack of direction</td><td>• Several Connections distinguish this level
• Dynamic Problem Solving Events</td></tr>
<tr><td rowspan="2">4</td><td>Abstract</td><td>NA</td><td colspan="2">Abstract concepts and procedures into context of alternative mathematical systems without concrete models</td></tr>
<tr><td></td><td colspan="3">• Not practical in secondary classrooms or textbooks Required to encompass entire mathematical continuum</td></tr>
</table>