

Feedback in Software and a Desktop Manufacturing Context
for Learning Estimation Strategies in Middle School

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Peter Malcolm

May 2013

© Copyright by

Peter Malcolm

All Rights Reserved

May 2013

Abstract

by

Peter Malcolm

Doctor of Philosophy in Education

Robert Q. Berry, III, Ph.D., Chair

The ability and to make good estimates is essential, as is the ability to assess the reasonableness of estimates. These abilities are becoming increasingly important as digital technologies transform the ways in which people work. To estimate is to provide an approximation to a problem that is mathematical in nature, and the ability to estimate is closely related to several key mathematical skills and aptitudes. These combined aptitudes allow individuals to judge reasonableness, or the proximity of estimates to the actual exact answer. Types of estimation such as computational and measurement estimation hinge on this sense of the reasonableness of an answer, but the concept of reasonableness itself is not given sufficient attention in primary and secondary grades in the U.S.

Digital technology provides the means to explore and refine student estimation strategies. Four studies were conducted with upper-elementary and middle school student participants. These featured the Estimation Calculator that provides immediate feedback on the percent error of estimates.

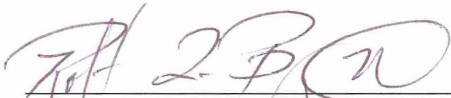
The first three exploratory studies provided a backdrop for the fourth study, which was conducted at a public middle school. This study also featured emergent desktop manufacturing technology. Such technology allows digital designs to be printed in three dimensions. The geometric nature of 3D design was used in conjunction with estimation instruction for a unit on computational and volume estimation. The guiding question in the research was whether instruction and practice with estimation-feedback software and 3D design software would influence students' estimates and strategies.

Estimation skills vary widely by class assignment. Students in a standard seventh grade mathematics class made estimates that were significantly less reasonable (*partial* $\omega^2 = .02$) than those of two advanced classes. The complexity of their strategies, as measured by digits that were included or ignored during rounding, was notably higher than their peers, suggesting that they were unfamiliar with simplification in estimation (Lefevre, Greenham & Waheed, 1993). An interaction effect (*partial* $\omega^2 = .03$) was present between the complexity of their estimates and those of their peers in advanced mathematics classes, suggesting an aptitude-treatment-interaction. Estimates using volume were hindered by computational estimation errors for all groups.


Curriculum, Instruction, and Special Education
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, Feedback in Software and a Desktop Manufacturing Context for Learning Estimation Strategies in Middle School, has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.



Robert Q. Berry III, Ph.D., Chair



Glen L. Bull, Ph.D., Co-Chair and Program Advisor



Jennifer L. Chiu, Ph.D., Committee Member



J. Patrick Meyer, Ph.D., Dean's Representative

March 19, 2013

Date

Dedication

To Lena, for your wisdom, patience and love

I could not have done this without you

Acknowledgements

Many thanks to Lena, for support and great suggestions. I am grateful to Matthew Reames, Willy Kjellstrom and Crystal DeJaeger, for helping collect and code data and improve the estimation model. I would like to acknowledge Dr. Mabel Kinzie for fine-tuning the research plan for Study 2, and Dr. Gail Hunger, for collaborating on the Estimation Calculator research from the very beginning. Thanks also to Eddie Pan, for helping put brand new code on WISEngineering, and to my mother, Sandra, and Kris Wiley, for helping sharpen the focus. Finally, thank you to the collaborating teachers: Mike Remchuk, Cheryl Knight, Jeannie Ballard, John Fisher, Tom Trotter, Anne Deane, Carrie Mersch & Darcy Sanderson.

Table of Contents

Dedication	iv
Acknowledgements	v
List of Tables	ix
List of Figures	x
Chapter 1	1
Introduction	1
Research Questions	5
Hypotheses	8
Logic Model	10
Chapter 2	16
Frameworks	16
Estimation	24
Feedback in Digital Technology	43
Conclusion	48
Chapter 3	50
The Estimation Calculator	50
Study 1: Exploratory Research with the Estimation Calculator	51
Study 2: Refining the Estimation Process Model	62
Study 3: Volume Estimation	69

Study 4: Estimation in Context	6
Chapter 4: Results	22
Introduction.....	22
Reasonableness of Estimates Across All Groups	26
Simplicity and Nuance for the Estimation Groups	35
Simplicity and Nuance for the Exact Calculation Group.....	51
Theme: Estimating the Volume of a Cylinder	55
Results of Student Interviews	56
Snapshots from the Intervention: Students' 3D Products.....	73
Results from a Follow-up Interview with the Teacher	75
Outliers.....	82
Conclusion: Standard Group Student Difficulties	83
Chapter 5: Discussion	85
Introduction: Is Estimation Teachable?	85
Summary of Results	87
Misuse of Strategies	97
Implications: Relevance of the Findings.....	98
Extending Prior Research	104
Limitations	107
Directions for Future Research	108

Conclusion	111
References	114
Appendix A: Problem Set for Study 4	123
Appendix B: Problem Set for the Interview Segment of Study 4.....	130

List of Tables

Table 1: Item Analysis for Interview	66
Table 2: Correlations Among Spatial and Computational Estimates	4
Table 3: Data Collection Strategy	9
Table 4: Activities for the Treatment Groups	15
Table 5: Medians and IQR for CE Reasonableness	28
Table 6: Descriptives for Reasonableness	29
Table 7: ANOVA for CE Reasonableness	30
Table 8: Medians and IQR for VE Reasonableness	32
Table 9: ANOVA for VE Reasonableness	33
Table 10: Overall Response Rate	36
Table 11: Descriptives for CE and VE Problems	37
Table 12: Medians and IQR for CE Complexity	38
Table 13: ANOVA for CE Complexity	39
Table 14: Complexity for Raw Number Problems	40
Table 15: Overall Strategy Response Rate	41
Table 16: Medians and IQR for VE Complexity	42
Table 17: ANOVA for VE Complexity	43
Table 18: Complexity for Geometry Problems	43
Table 19: Algorithmic Solutions	45
Table 20: The Nuance Scoring System	46
Table 21: Nuance Scores	47
Table 22: Medians and IQR for Nuance	48

Table 23: ANOVA for Nuance	49
Table 24: Complexity Scores for CE	52
Table 25: Complexity Scores for VE.....	53
Table 26: Nuance Scores for LAX Group	53
Table 27: Within Group Interview Observations.....	67
Table 28: Interview Estimate MAMEs	69
Table 29: Summary of Results.....	86

List of Figures

Figure 1: Estimation Logic Model.....	9
Figure 2: Process Model	29
Figure 3: Estimation Error Bar	50
Figure 4: Initial Original Process Model	52
Figure 5: Remote Tutoring Link	53
Figure 6: Revised Process Model	55
Figure 7: Final Process Model	67
Figure 8: The Spatial Estimator	71
Figure 9: The CE Assessment Interface.....	84
Figure 10: The VE Assessment Interface	85
Figure 11: Boxplot of CE Reasonableness	101
Figure 12: Marginal Means CE MAME	104
Figure 13: Boxplot of VE Reasonableness	105
Figure 14: Marginal Means VE MAME.....	107

Figure 15: Boxplot of CE Complexity.....	111
Figure 16: Boxplot of VE Complexity.....	115
Figure 17: Boxplot of Nuance.....	122
Figure 18: Marginal Means Nuance	124
Figure 19: Estimate Errors COE0401	132
Figure 20: Estimate Errors COE0402	134
Figure 21: Estimate Errors SAE0317	136
Figure 22: Estimate Errors SAE0306	137
Figure 23: Estimate Errors LAX0702.....	139
Figure 24: Estimate Errors LAX0725.....	140

Chapter 1

Introduction

The Importance of Assessing Reasonableness in a Digital Age

Over the last half century, the growing presence of information technology has steadily transformed the type of work that people have been required to do (Autor, Levy & Murnane, 2003). Simple, repetitive and rule-based jobs that were formerly done by hand are increasingly automated, while there is a growing need for persons capable of "performing nonroutine problem-solving and complex communications tasks" (Autor et al., 2003. p. 1279). One such nonroutine task is estimation, and the research described in this paper concerns estimation in two forms. Before formally defining computational and measurement estimation, it is worth discussing the relationship between estimation and technology.

Digital devices that can help compute solutions more quickly and easily than people can have not rendered mental computation skills an "anachronism" (Hope, 1986). Primary and secondary students are taught to perform calculations by hand. Mathematics educators seek to instill a sense of the reasonableness of numbers in context (Garofalo & Bryant, 1992; Rheinlander et al., 2008). This sense of reasonableness allows students to identify and fix typographical or technological errors, such as wrong answers on a calculator (Berry, 1998).

The recent financial crises caused by rapid devaluation of dot-com companies and U.S. real estate had complex underlying causes. As Schiller (2004) suggests, both crises were fueled by "irrational exuberance" and misinterpretation of data. As an economist

who predicted both crises by rationally analyzing the underlying data, Schiller's methods address what mathematics educators would call the reasonableness of answers. When the "answers" are statistics on volatility in multi-trillion-dollar international financial markets, reasonableness matters. Speculative bubbles predate the advent of information technology of course, but technology accelerates the operations of the financial world and can even contribute to volatility (Easley, López de Prado & O'Hara, 2010). It is safe to say that in 2012, there is no computer-based substitute for reasonableness.

This dissertation describes research that addresses this concern. Perhaps computers can support people in learning to do what the computers cannot, which is to assess the reasonableness of answers and thereby make good estimates. If so, this technology will aid mathematics teachers and provide learners with estimation abilities that matter both professionally and in informal contexts.

The various forms of estimation have innumerable practical applications in everyday society (Usiskin, 1986). Professionals such as architects provide estimates, usually in terms of time and financial cost, when bidding on a contract to complete a large job. Less consequential opportunities for estimation occur in an adult's life every day, such as the need to estimate grocery costs or travel times.

The National Council of Teachers of Mathematics emphasizes the importance of estimation abilities (NCTM, 2000, 2006) such as the ability to make reasonable computational estimates. For teachers, estimation provides a way to teach students flexibility with mathematics (Siegler & Booth, 2005). Computational estimation ability is a strong predictor of success in the SAT-Mathematics test (Hanson & Hogan, 2000; Hogan, Wyckoff, Krebs, Jones, & Fitzgerald, 2004; LeFevre, Greenham & Waheed,

1993). There is reason to believe that this is at least partially a causal link because efficient estimation permits rapid elimination of distractors on such a multiple choice test. It has been linked to general math ability (Levine, 1982) and tied to specific skills such as working with place value and multiples of ten (Rubenstein, 1985).

Computational estimation ability is closely related to an understanding the reasonableness of an answer (Rubenstein, 1984; Dowker, 1992; NCTM, 2000). A recent study based on focus groups with forty U.S. teachers (Rheinlander et al., 2008) indicates that reasonableness is still an underutilized and concept in the K-12 classroom and one that requires more explicit instruction.

Reasons for Estimating

The 1986 Yearbook of the NCTM (Schoen & Zweng, 1986) was devoted to estimation and mental computation. Bolstering the argument for the importance of estimation, Usiskin (1986) listed four reasons for its use in everyday situations:

1. constraints force estimates
2. estimates increase clarity
3. estimates are easier to use
4. estimating gives consistency (pp. 3-11)

Usiskin emphasizes the first reason, the existence of constraints, as the most common reason for estimation in everyday practice. Estimation is necessary when actual values are unknown or unknowable, such as predictions of future events or generalizations to populations from samples. The second reason, clarity, follows from the fact that most people have short-term memory that can handle only a few digits. Short-term memory is referred to more precisely as working memory (Baddeley, 1986)

and is one of the central concerns of research on computational estimation. Working memory is discussed in more depth below.

Usiskin's third reason, ease of use, is closely tied to this concern of clarity. Both clarity and ease of use vary directly with economy of representation (Bruner, 1966). Whether one is performing calculations by hand or mentally, the process of computation is simpler and thus easier with numbers having fewer significant digits. Finally, Usiskin argues that estimates are used for consistency, citing conventions such as reporting of sports scores, unemployment figures, and world records for racing. Each convention is partially rooted in tradition and the practical validity of the estimated value. One overarching consideration is that estimates that are themselves based on other estimates cannot meaningfully contain more significant digits than the original estimates. For example, if a car used approximately 8.5 gallons of gas to travel 240.2 miles, it does not make sense to report the resulting estimated miles per gallon (mpg) of the car as 28.258824 . Instead, the mpg should be rounded to 28 so that it does not contain more significant digits than the two from the 8.5 gallon estimate.

Multi-step Geometry Status in U.S. Education

Although the U.S. attained competitive mathematics scores on the 2007 TIMSS study, geometry scores were notably below-average. Multi-step geometry in particular proves surprisingly difficult for students, with students performing worse on multi-step problems than their separate performance on each individual step would predict (Kao, Roll & Koedinger, 2007). The process of estimation is designed to simplify mental computation (LeFevre et al., 1993; Usiskin, 1986). Volume estimates may be computed based on estimates for dimensions of a shape. On interesting outcome of this research

will be the result of the interaction between the subgoaling difficulties of multi-step geometry problems (Kao et al., 2007) and the simplifications that an estimator may introduce to reduce cognitive load (Sweller, van Merriënboer, & Paas, 1998).

Estimation and Digital Fabrication

Multiple national organizations have emphasized the value of technology to enhance mathematics teaching. The NCTM (2000) highlights the importance of the use of technology for learning, and the strategic use of technology tools is one of the Common Core State Standards for Mathematical Practice (National Governors Association, 2010b). The digital fabrication (DF) and desktop manufacturing work described in Study 4 deals with specific mathematical knowledge in what is believed to be a motivating context (Berry et al., 2010). Testing the hypothesis that DF provides affective motivation is beyond the scope of the current research however. Here DF will be provided as a framing context and a rationale for students' estimations. Generating estimates before starting a project help to establish time and cost parameters in engineering, and the kinds of multi-step estimates under investigation here have been incorporated into engineering education (e.g., Dunn-Rankin, 2001). There are natural connections between engineering as practiced by professionals and the K12 mathematics curriculum that are often unexplored by K12 educators (Katehi, Pearson & Feder, 2009) and this work seeks to draw upon and exploit those connections (p. 156).

Research Questions

A significant body of research exists on computational estimation. This includes studies on the mathematical processes of computational estimation (Levine, 1982; Reys, Rybolt, Bestgen & Wyatt, 1982), the skills needed to perform computational estimation

(Rubenstein, 1985) and the corresponding cognitive processes involved (LeFevre et al., 1993; Lemaire & Lecacheur, 2011; Siegler & Booth, 2005; Tversky & Kahneman, 1974).

Non-empirical literature suggests tools and techniques for teaching computational estimation (Rose & Morgan, 1986; Wheatley & Hersberger, 1986). Very few studies have been done to investigate how well these kinds of tools actually work (e.g., Fung & Latulippe, 2010; Lan, Sung, Tan, Lin & Chang, 2010; Yang & Hsu, 2009). Thus the driving question behind all of the research in this paper is, *Is estimation teachable?*

While this question is too broad to be answered by a single study, the research described here attempts to address key elements of estimation and the ways in which addressing those elements might influence student performance. A question that falls under this driving question is, *Is a familiarity with context sufficient for making reasonable estimates for problems in that context?*

Computational estimation is applicable in many contexts, and the research described here places computational estimation within a context of volume estimation for DF. Volume estimation can be performed either as a series of mental calculations based on the dimensions of a solid shape, or by comparing a shape in its entirety to other familiar shapes, which can serve as reference points (Joram, Subrahmanyam & Gelman, 1998). Estimation is characterized by a tension between simplification and proximity (LeFevre et al., 1993). Estimating the volume of a shape by comparing it to another shape might be a simpler strategy than one involving mental computation, however it might also result in less proximal (i.e., less reasonable) estimates. All three groups in Study 4 were given extensive exposure to the context of multi-step geometry problem

solving. Their resulting familiarity with the context was reflected in measurements of reasonableness (error) in their post-assessment estimates.

RQ1: Can instruction on computational and volume estimation combined with automated feedback on estimation in a project-based context lead to students making:

- a) more reasonable computational and volume estimates?
- b) simpler computational and volume estimates?
- c) more nuanced computational estimates?

RQ2: Can an identical project-based context that teaches exact volume computation rather than estimation, and does not include automated feedback lead to students making:

- a) more reasonable computational and volume estimates?
- b) simpler computational and volume estimates?
- c) more nuanced computational estimates?

R3) How does reasonableness of estimates differ between students after the two treatments described in R1 and R2 for:

- a) computational estimation problems?
- b) volume estimation tasks that resemble the problems in the unit?
- c) multi-step volume estimation problems that do not resemble the problems in the unit?

RQ4: How do students in the two conditions describe their computational and volume estimation strategies:

- a) before a project-based unit?
- b) after a project-based unit?

Hypotheses

It was hypothesized that estimate error feedback would improve computational estimation ability but would actually hinder improvement on volume estimation ability. This was based on the premise that computational ability would benefit from quick, repeated practice since it is a skill that middle school students of this age are capable of (Case & Sowder, 1990) but for which they lack automaticity (Neves & Anderson, 1980). The multi-step geometry task, on the other hand, was hypothesized to elicit gaming behavior from students (Baker et al., 2009). In other words, students were anticipated to use this software as a crutch, because students would deem random guessing with feedback to be less effortful than revising estimates.

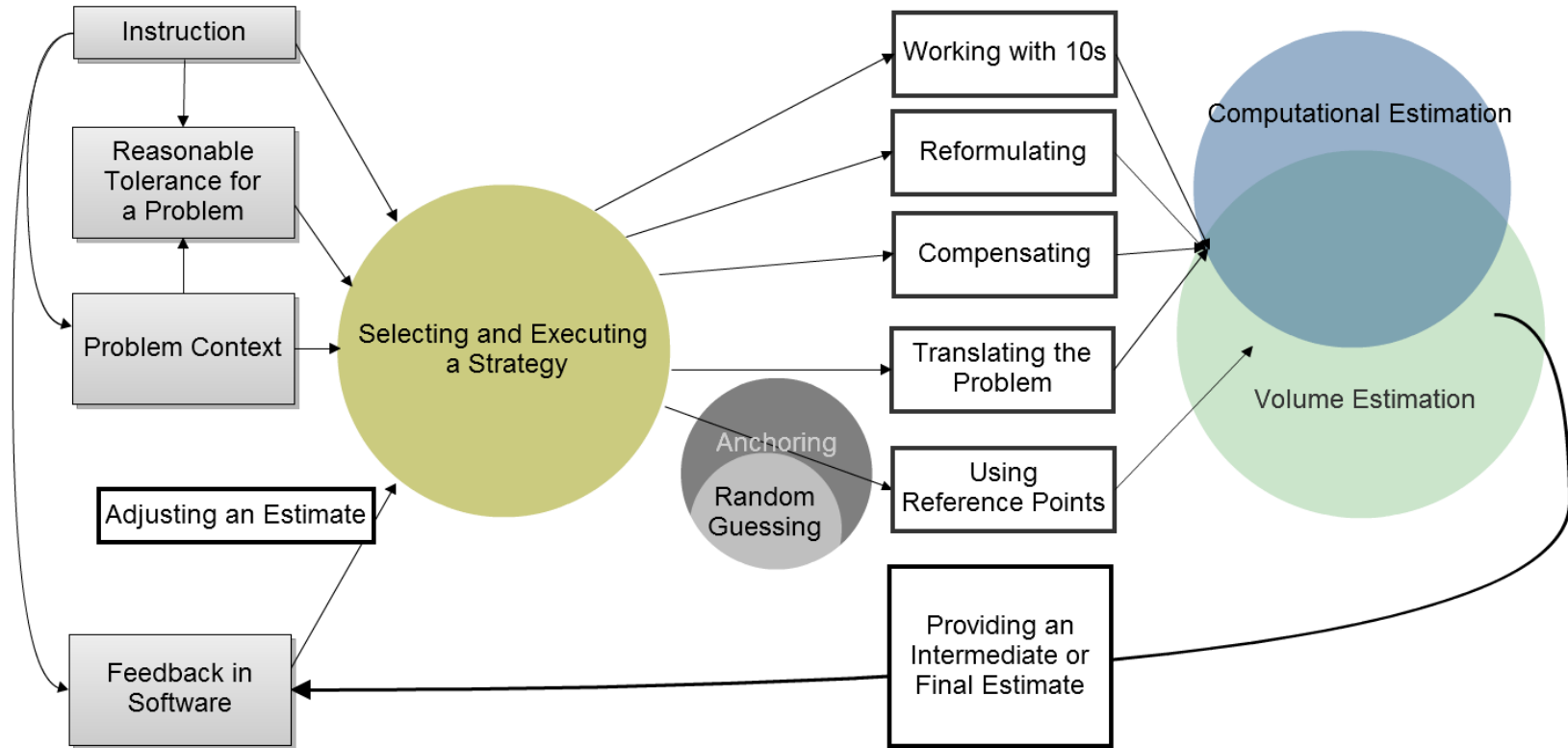


Figure 1: Estimation Logic Model

KEY: Dark Rectangles – Independent Variables
 Light Rectangles – Dependent Variables (Observed)

Circles – Dependent Variables (Inferred)
 Arrows – Causal Links

Logic Model

Independent Variables

The logic model underlying this research (see Figure 1) starts with instruction. Instruction is used in a broad sense here to refer to the entire instructional design including all whole-classroom lectures, individual or small group activities and any supporting technology. Any activity on the topic that the teachers, students or the supporting devices do is considered instruction. Three components of instruction are important in this logic model:

- 1) Classroom discussion on estimation processes and strategies
- 2) Problem Context
- 3) Feedback

In the diagram (Figure 1), a causal arrow leads directly from the independent variable of instruction to Selecting and Executing a strategy. Classroom discussion is expected to influence estimation strategy selection and execution. A teacher can directly teach students how to work with powers of 10, how and when to round, etc. The dependent variables (white rectangles) at the middle of the model are all abilities that students are to learn to become good estimators, and all of these can be taught directly.

The term tolerance is borrowed from the engineering lexicon to mean "the allowable deviation from a standard" (Merriam-Webster, 2012). The reasonable tolerance for an estimation problem depends upon its context. Some contexts, such as grocery shopping, have fairly large tolerances. The tolerance can also be asymmetrical. For example, an overestimate on a financial cost may be "erring on the side of caution," while an underestimate may be more risky. Estimates for cost are also sensitive to

available funds. All of these considerations influence estimation strategy selection and execution. In the research here, a tolerance of $\pm 15\%$ has been chosen for computational estimation problems without context. In problems with a specific context, the tolerance can be a topic of class discussion and can be chosen via classroom consensus. Because it may be chosen by instructional designers, and because it strongly influences the processes of estimation, the tolerance is an independent variable in the research. Problem context is another, interrelated independent variable.

Feedback is the final independent variable. Much of the research described here involves interfaces that provide feedback on estimation. The form and timing of the feedback influence how students adjust estimates. It is posited in the logic model (Figure 1) that students learn while make these adjustments, and that this learning impacts their abilities to select and execute strategies. Feedback is information that allows students to confirm, disconfirm or "tune" their knowledge (Butler & Winne, 1995). It is defined more explicitly in the next chapter.

Variable Manipulation Treatments

Research Questions 1 and 2 vary the instruction and feedback variables, while keeping the context variable constant. Three groups of students received two treatments in Study 4. These were the estimation and the exact answer treatment, respectively. Two groups of students received the estimation treatment and the third group of students received the exact answer treatment. (The reasonable tolerance variable is not relevant for the exact group.) Both treatment groups received instruction on desktop manufacturing and computer-aided design, and both focused heavily on the mathematics of making volume estimates. This provided the context for the unit. It included structure

and goals of the desktop manufacturing activity, including requisite estimates of volume and materials needed to fabricate designs.

Over 40% of the variance in computational estimation ability is attributable to a person's ability to work with place value and powers of 10 (Rubenstein, 1985). The same study showed that a tiny percentage (about 1%) of variance in estimation ability is attributable to how well an estimator gets to know the problem. The research described in Study 4 compared three groups of seventh-grade students. This shed some light on the remaining variance in computational estimation ability. The results suggest that, for a brief intervention such as was used in the present study, prior mathematics ability will have a more significant influence on estimation ability than instruction and practice combined. However, for generating volume estimates without the use of any computation, anchoring seems to support students in making reasonable estimates regardless of mathematics ability. The term anchoring is explained in the subsequent section.

Dependent Variables (Inferred)

Tversky and Kahneman (1974) use the term anchoring to describe the starting point for estimation. The magnitude of a numerical anchor has a profound impact on a person's final estimate. If a researcher (or in the case of the current study, a teacher) suggests a number as a starting place for estimation, the person making the estimate (research participant or student) is likely to make a final estimate that is close to that anchor. The technique for measurement estimation described by Joram et al. (1998) depends upon the use of a physical object as an anchor. They refer to this as a reference point. The repeated practice within a certain context proscribed by this research will

provide both treatments with a ready supply of mental anchors as reference points for volume estimation problems on a post-assessment task. This may improve volume estimation on post-assessment problems for both groups. Anchoring is one of the inferred dependent variables in the logic model, and it is directly related to the observable variable of the use of reference points.

Spurious anchors were the independent variables in Tversky and Kahneman's (1974) studies. The research described in Study 4 avoided this. Only logical anchors, if any, were provided to the estimation groups during practice or to either treatment group during pre- or post-assessment. However, a previous study (chapter three, Study 3) revealed that students are liable to choose their own illogical anchors for computational estimation problems. The distinction that was drawn in Study 3 between informed guessing and random guessing (See Figure 7) is equivalent to the distinction between the use of logical anchors and illogical anchors. Logical anchors are likely to result in efficient strategies with the Estimation Calculator and reasonable estimates. Illogical anchors are likely to result in floundering and are unlikely to converge on reasonable estimates in an efficient manner. This definition rests on the assumption that even seemingly random guesses have some cognitive basis. Hence random guessing is considered a special case of anchoring where the chosen anchor is illogical for the problem.

The overlapping constructs of computational estimation and volume estimation include regions in the logic model separate from anchoring and random guessing. These are both considered reference-point-based strategies. Computational estimates can be

made for problems that do not pertain to volume, and volume estimates can be made without using any numerical computation.

Dependent Variables (Observed)

The observable dependent variables in this study are shown at the center of the logic model. The five observable variables that relate to estimation are:

- 1) Working with 10s
- 2) Reformulating
- 3) Compensating
- 4) Translating the Problem
- 5) Using Reference Points

Working with powers of 10 is a standard component of elementary mathematics education (e.g., National Governor's Association, 2010b) and contributes over 40% of variance to estimation ability (Rubenstein, 1985). The following three items refer to processes – which encompass specific strategies – for estimation. They are described in more detail in the following chapter. Reformulation is a process that includes rounding strategies. Rounding has been standard elementary educational practice for many years, and algorithmic rounding is often the only computational estimation strategy that children are taught in school (Reys & Bestgen, 1981). Compensation is a term coined by Reys et al. (1982) to refer to the process of adjusting operands or an intermediate estimate to compensate for having rounded. Translation is another process described by Reys et al. (1982) that consists of altering the structure of a problem, such as converting repeated addition to multiplication. Finally, the use of reference points is a process helps define

measurement estimation (Joram et al., 1998). More broadly, a reference point may be a numerical anchor (Tversky & Kahneman, 1974).

Each one of these five dependent variables is measurable using interview data. If computer activity elicits intermediate solutions – asking students to show their work – then these estimation processes can be directly detected or inferred. For example, if a student is asked to make an estimate for 62×23 and enters 60×20 as an intermediate solution, it can be inferred that a student has reformulated the problem via rounding down.

The most important dependent variable is an estimate itself, shown at the bottom of Figure 1. This may be an intermediate estimate, such as 120 as an estimate for 60×20 , which a student may revise after he or she receives feedback. A series of estimates for the same problem can reveal patterns in student estimation behavior (see Studies 1 and 2 in the third chapter).

Finally, there is the dependent variable of estimate adjustment, at the bottom-left of Figure 1. Adjustment is measurable either as the quantifiable difference between two sequential estimates (e.g., 120, then 1200), the qualitative difference between two sequential reformulations or translations of a problem (e.g., 60×20 , then 600×2), or by student description ("I added the zero back in."). Estimate adjustment is closely related to the other dependent variables listed above.

Chapter 2

The following review presents published literature and unpublished literature that was gathered using multiple methods. First, faculty and mentors were asked for suggestions on particular publications or researchers whose work relates to the present study. Faculty and mentors provided direct access to published work, and in one case, to a relevant unpublished master's thesis. Works cited in these publications were sought out and included if they were relevant and sufficiently rigorous. Online databases, including EbscoHost, JSTOR, ERIC, and Google Scholar were also used extensively. The library database at the University of Virginia also provided valuable sources of literature. Search terms included "estimation" and "feedback," and combinations such as "computer-based feedback" were used as well.

The next four sections describe four frameworks that are foundational to this research. These are, the KLI framework (Koedinger, Corbett & Perfetti, 2012), cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller et al., 1998), automaticity, and digital fabrication (Berry et al., 2010; Bull & Groves, 2009; Gershenfeld 2005).

Frameworks

The following frameworks draw heavily upon cognitive psychology. Snow and Lohman (1993) suggest that cognitive psychology can contribute to educational measurement by improving construct validity, providing diagnostic information, and expanding the research agenda. Here the use of cognitive approaches for diagnosis, or formative assessment, is of primary interest. The research described here involves the use of diagnostic tools that seek to address unseen, or latent constructs of understanding.

These constructs were broken down, a priori, for individual analysis as knowledge components (Koedinger, Corbett & Perfetti, 2012). The conclusions that can be drawn from such diagnostic tools with binary response data are limited (Norris, Macnab, Phillips & 2007) in that no conclusions whatever can be drawn about motivation. Moreover, the research here concerns not just *if* an estimate is unreasonable but *how* unreasonable it may be. The research described in the present studies provides richer sources of data, including students' descriptions of their strategies and actual error percentages for their estimates.

The incorporation of desktop manufacturing is also intended to provide students with anchors for volume estimation, as well as motivation for the activities students performed. The methods bear some resemblance to problem- and project-based learning frameworks (Blumenfeld et al., 1991), which have been found to motivate students by contextualizing learning.

The KLI Framework

The knowledge-learning-instruction (KLI) framework (Koedinger et al., 2012) is based on the premise that all knowledge can be learned and assessed in elemental chunks called knowledge components (KCs). At their smallest grain size, KCs can be equivalent to production rules, which are if-then statements that are used to perform actions.

According to the ACT-R theory of learning (Anderson, 1993), people learn procedural information in the form of logical rules. Procedural knowledge refers to action knowledge, or the ability to do things (Snow & Lohman, 1993). Here it is worth noting that KCs can also consist of declarative knowledge, which is factual knowledge.

Declarative knowledge informs procedural knowledge but does not map directly to

action. Production rules on the other hand, map situations directly to the actions that should be taken in those situations. These can be formally written as if-then statements, which can be written in machine-readable form (Anderson, 1993; Anderson, Boyle, Corbett & Lewis, 1990; Anderson & Pelletier, 1991).

The framework matches learning processes to KC types. The term KC types refers to general classes of KCs: facts, rules or principles. The learning of rules (a type of KC) is most appropriately matched to an induction and refinement learning process, while the learning of facts can be depend upon a memorization learning process, etc. However there is an asymmetry; the cells below the diagonal represent learning processes (rows) that are simpler than the intersecting KC types (columns). The plus signs (+) in these cells signify that there is some value in learning processes that are simpler than the KCs being taught. Thus, while rule learning is best matched to induction and refinement, memorization plays a supporting role in rule learning.

The present research pertains to digital fabrication and estimation. Desktop manufacturing is the specific type of digital fabrication under investigation in this research. Digital fabrication is the creation of physical objects from digital designs in a learning context (Bell et al., 2010; Berry et al., 2010). Personal fabrication (Bull & Groves, 2009; Gershenfeld, 2005) is the term generally used for the DF procedure when it is used in homes and informal contexts, and desktop manufacturing involves hardware that is small enough to fit on a desk in a home or office. Digital fabrication is certainly not a single KC, since it requires an array of technical skills from software knowledge for computer aided design (CAD) and hardware knowledge for computer aided manufacturing (CAM) to the initial creativity of product design. One of the clearest

connections between this set of skills and the most current national standards for learning in the U.S. (National Governors Association, 2010a) is in the use of solid geometry for CAD. Currently, simple CAD tools exist that are appropriate for student use but which do not provide a software-based method for finding the volume of a three-dimensional design. In an educational context this can be an asset. For the research described in Study 4, students were asked to make volume estimates, but could not find the volume in the design in the software as they did generated their designs.

While estimation might be considered to be a KC in its own right, it is certainly an integrative KC (Stamper & Koedinger, 2011) composed of smaller KCs such as the abilities to round and perform basic arithmetic operations. The following sections describe in greater detail how estimation might be decomposed into individual KCs.

Cognitive Load Theory

Cognitive load theory (Paas, Renkl, & Sweller, 2003; Sweller et al., 1998) posits that all instruction imposes a cognitive load on the working memory of the learner. During problem solving, this load can consist of extraneous or germane load. The germane load is that which is useful or important, while the extraneous load results from problem features that are not directly relevant to its solution. The theory suggests that all problems have an intrinsic cognitive load, stemming from the inherent complexity of the problem. Element interactivity is one of the central metrics of intrinsic cognitive load. If elements of a problem cannot be treated and learned separately, then that problem's intrinsic load is higher than a problem that can be taught and learned in a more piecemeal fashion. Taken in general, estimation can be considered to have a relatively low intrinsic cognitive load. For example, mentally managing front-end digits and managing place

value can be treated separately. However, certain estimation problems have high element interactivity because they involve large numbers of digits and challenging mathematical components such as carrying in addition (Kalamian & LeFevre, 2007).

Learning as Automaticity

Kahneman (2011) describes human mental functioning metaphorically as consisting of two systems, the automatic system and the effortful system, which he labels System 1 and System 2 respectively. This distinction is based on his years of psychological research on the conflicts that often occur between unconsidered and considered responses to judgment or problem-solving questions. Recent behavioral and neuropsychological evidence supports the existence of two such systems for processing numerical information (Feigenson, Dehaene & Spelke, 2004).

Automaticity is measurable as a reduction in errors, which is quantifiable and has been found to follow a decreasing power function (Newell & Rosenbloom, 1980). In short, more practice leads to fewer errors. The process of automating knowledge has been compared to the process of compilation of software (Neves & Anderson, 1980). It occurs as a shift from deliberate and effortful behavior to one that occurs without direct thought. Throughout the present research, the construct of learning is considered to be essentially one of increasing automaticity. The goal of the present research, which is to teach estimation ability, is more specifically to allow students to become more automatic.

Instruction on estimation strategies is an independent variable in the present study. Instruction involves presenting information to a student in such a way that it can be learned. The term "direct instruction" will be used here in approximately the way it is meant by Kirschner, Sweller, and Clark (2006). Their definition exists largely in contrast

to a cluster of alternative instruction theories, most notably discovery and experiential learning. While the philosophy espoused in the present research does not oppose inquiry learning, it is useful to refer to Kirschner et al.'s (2006) theory of teaching, and use the term "direct instruction" as shorthand for simply telling students what they need to know.

Initially, however, a novice needs a relevant context for information. A level of expertise, or automaticity in performance, will never be attained if a novice never establishes a foothold in a content area. The next section discusses context, which fosters the early stages of learning.

Desktop Manufacturing as a Context

Bransford, Brown and Cocking (2000) present a strong case for presenting information in a context. Students are unlikely to retain information if it is presented devoid of context. This principle is also central to learning frameworks as knowledge integration (Linn, Lee, Tinker Husic & Chiu, 2006) that posit that knowledge should be presented so that it connects both to students' current knowledge and to extend it in ways whose relevance students can easily perceive. Fabrication, as envisioned by Gershenfeld (2005) provides individuals an opportunity to design and create objects that they need or find interesting. He describes it as the ability to "make (almost) anything," which is an idea that is instantly approachable to students, and can capture their imaginations.

Digital Fabrication and Desktop Manufacturing

Technologists and technology researchers (Bell et al., 2010; Gershenfeld 2005) foresee personal fabrication as an oncoming cultural upheaval similar to the one brought on by personal computing. They predict that, just as computers evolved from many-ton devices that occupied multiple rooms, so will the large manufacturing machines of our

modern age evolve to smaller, cheaper, and more portable devices. This has already occurred for die-cutters and additive printers for plastic (Bull & Groves, 2009). The advent of these technologies has the potential to empower individuals to develop everyday objects in their own homes (Gershenfeld, 2005). The possibilities of this for classroom and preservice education research have been proposed and tentatively explored with promising results (e.g., Berry et al., 2010; Tillman, 2011; Tillman et al., 2012). For the purposes of the present research, digital fabrication is "the process of translating a digital design developed on a computer into a physical object" (Berry et al., 2010, p. 2). Desktop manufacturing is a subset of digital fabrication that involves 3D printers that are small enough to fit on a desktop in a classroom or office.

Project-Based Learning

Thomas' (2000) review of project-based learning finds that conflicting definitions exist for this term. He draws heavily upon early work in the area by Blumenfeld et al. (1991) and distills five criteria by which to judge whether or not a curriculum can or cannot truly be termed project-based. For Study 4 of the present research, which is described in chapters 3 and 4 of this work, the most relevant definition concerns the incorporation of a project that is, as Thomas (2000) describes, "central, not peripheral to the curriculum." The desktop manufacturing component of Study 4 is a project in which students design pencil-holders.

Project-based has well-known precedents in science education (e.g., Blumenfeld, Fishman, Krajcik, Marx, & Soloway, 2010). While she does not refer to them as "project-based," studies described by Boaler (2000) embody the principles of project-based learning in the context of mathematics. Project-based science and project-based

mathematics are designed to have a strong motivational component, and often involve the use of technology. In fact, Blumenfeld et al. (1991) list benefits of using technology in project-based instruction, one of which is that technology can diagnose and correct errors. The technology tools described herein are designed to help students diagnose and correct their own errors when performing estimation.

Technology in the Classroom

Technology has come to be shorthand for digital technology, and as such has been a feature in the field of education for the last half-century (Saettler, 1990). New digital technologies often draw upon the traditions of intelligent tutoring systems (ITS) and more broadly, computer aided instruction systems (CAI). The chapter section entitled Feedback in Digital Technology, below, describes Cognitive Tutors (Anderson et al., 1995), which are among the most sophisticated manifestations of this line of work to date. These represent a highly structured approach to teaching for what Jonassen (2003) would call learning *from* computers. In Jonassen's view, a more fruitful approach is to lead students to learn *with* computers. He describes a philosophy, based on constructivism, wherein students use computers to build with, build from, and refine what they know. Broadly speaking, there has been a slow adoption from more exploratory, constructivist uses of computers in K-12 school that has followed the early adoption of more didactic approach. Seymour Papert (2005) draws an extended analogy between the U.S. educational system's adoption of technology and Piaget's theory of learning. He claims that the adoption of didactic, CAI-style tools is the educational system assimilating digital technology into a pre-existing schema. Subsequently education in the U.S. has accommodated more exploratory and open-ended uses of technology

Papert (1980) envisions computers as tools with which students can explore phenomena. He uses the term computation not to refer to numerical computation but rather to refer to the writing and execution of computer programs. By extension, he sees children developing a sense of the formalisms and structures of mathematics through the process of creation. The desktop manufacturing used in Study 4 is intended to capture some of the creative mathematical spirit about which Papert wrote. Students in Study 4 developed their own pencil holders, while adhering to real-world constraints. The uses of the Estimation Calculator described in Studies 1, 2 and 4, and in a modified form in Study 3 allow for an open-ended approach to mathematical problem solving that does not prescribe a solution path. Arguably, when one is using one of these online estimation tools, any strategy that consistently results in reasonable estimates is a valid strategy for estimation.

Estimation

The following sections describe estimation. First, different kinds of estimation are distinguished and computational and measurement estimation are defined. After that, computational estimation is analyzed in more depth. Cognitive elements of estimation, and then pedagogical concerns around estimation are presented.

Distinguishing Computational from Measurement Estimation

Siegler and Booth (2005) provide an extremely broad definition of estimation as "a process of translating between alternative quantitative representations, at least one of which is inexact" (p. 198). It is worth noting that the ultimate representation is the estimate itself, which should be simpler than the prior representations and should be more inexact than they are. The definition used here considers an exact answer to be an

estimate, albeit one with 0% error. Thus it is possible for all representations used in the estimation process to be exact (LeFevre et al., 1993) although this violates the spirit of estimation. Siegler and Booth (2005) draw a sharp distinction between estimation that requires real-world knowledge and estimation that can be provided context-free. Their (2005) review of numerical estimation literature focuses on the former category, specifically computational, numerosity and number line estimation.

Hogan and Brezinski's (2003) study on undergraduates, using a battery of tests, included a principal component analysis on three kinds of estimation, as well as number facility and quantitative reasoning. The first type of estimation, computational estimation (CE) loaded very strongly with number facility and quantitative reasoning. Meanwhile, measurement estimation (ME) loaded along with numerical estimation (quickly judging the number of dots in an array) as an orthogonal component. Based on these results, the abilities of CE and ME can likely be considered to be independent abilities.

Computational Estimation

Computational estimation is defined in (Reys, Rybolt, Bestgen & Wyatt, 1982) as:
The interaction and/or combination of mental computation, number concepts, technical arithmetic skills including rounding, place value and less straightforward processes such as mental compensation that rapidly and consistently result in answers that are reasonably close to a correctly computed result. This process is done internally, without the external use of a calculating or recording tool. (p. 6)

Similarly, the focus in much of the estimation literature is that an estimate must be a reasonable (Dowker, 1992: NCTM, 2000) or satisfactory (LeFevre et al., 1993)

approximation to an exact answer. For the research the above definition is used, although with a slight modification on the clause "without external use of a calculating or recording tool". The technological instruments used in this research provide feedback on the steps of estimation and can be used as a calculating or recording tool in some cases.

Measurement Estimation

Joram et al. (1998) define measurement estimation informally by way of examples, such as everyday estimates adults make of driving distances or the weight of produce at the grocery store. For the present study it is used to mean any real value in the world for which a person requires an estimate. This stands in contrast to computational estimation, which concerns numbers that are the result of mental operations on other numbers. Kahneman and Tversky (1974) found that the heuristic of anchoring plays a large role in real world estimation. This has been found to be true for measurement estimation as well as for computational estimation (LeBoeuf & Shafir, 2006).

The research that is described in Study 4 involved measurement estimation in an on-screen view. Research indicates that initial estimates of distance made in virtual environments depend on the orientation of the objects to be estimated (Richardson & Waller, 2005). When a single object is shown to a viewer from an egocentric perspective, that viewer will systematically underestimate how far away that object is (virtually) from his or her own body. In the case of Richardson and Waller (2005) this foreshortening effect was roughly -40% error on average. However, when two objects are shown in an orientation that is perpendicular to a viewer's line-of-sight, that viewer will tend to make a reasonable estimates (-10% error in Richard & Waller, 2005). One apparent remedy would be to allow users to rotate shapes so as to estimate the aspects,

such as length width and height, of three-dimensional shapes. If the user viewing the shape on screen can always orient the aspect at a 90° angle from his or her view, he or she is more likely to develop an accurate estimate. As a solution in Study 4, students were always explicitly told, either on-screen or on-paper, the dimensions of the shapes whose volume they were to estimate.

The remainder of this section concerns computational estimation, because the present research is designed to investigate the techniques and strategies used to form estimates for volumes in geometry. Beyond concerns of visual perspective and choice of an anchor, the development of volume estimates depends heavily on mental multiplication of dimensions of shapes. Once initial measurement estimates have been made, the rest of the cognitive work is in synthesizing and combining these estimates computationally.

Mathematical Abilities Related to Computational Estimation Ability

Rubenstein's (1985) regression analysis of computational estimation among eighth graders revealed the ability to operate with tens accounted for more than 40% of the variance in students' abilities to make reasonable computational estimates. The example problem she gives from the operating with tens section of her research instrument is a multiple choice question: $64.8 \div 10 =$

- a) 648 b) 64.80 c) 64.8 d) 6.48

(p. 110)

Another contributing ability was the ability to make comparisons between numbers, although this was only found to contribute 3% of additional variance in

estimate scores. Overall, Rubenstein (1985) did not find any regression that accounted for more than 47% of total variance.

Computational Estimation in Psychological and Psychometric Literature

Estimation ability measures can provide more information than instruments measuring other abilities because an estimate contains error, usually a relative error measured as the difference between a participant's estimate and actual value, divided by actual value. Other instruments such as multiple choice tests rely on binary measures, with each response marked as correct or incorrect. Estimation researchers (e.g., LeFevre et al., 1993; Levine, 1982) use relative error to rank and analyze participants' estimation ability. In other words, a research participant's estimation ability can be operationalized as standard or mean deviation of estimates from the actual solutions to estimation problems.

Cognitive Processes Involved in Computational Estimation

LeFevre et al. (1993) developed a model of computational estimation (see Figure 2) that includes encoding as the first step. Next, according to the model, the research participant attempts to retrieve the exact answer from long term memory. This model incorporates the terminology of Reys et al. (1982). In this model, individuals first seek an exact answer in chunked form from long term memory. They then begin the estimation process if the exact answer cannot be quickly retrieved or calculated. Key points to note about this model are that place value adjustment and post compensation are treated as separate steps. Also, the initial attempt to retrieve an exact answer may be repeated more than once on successive simplifications of the problem. Interestingly, adults will sometimes provide exact answers to problems even when they are explicitly

asked not to (Levine, 1982). This suggests that individuals' trained response to provide an exact answer will override given instructions in a computational estimation task.

From (LeFevre, Greenham & Waheed, 1993) p. 29

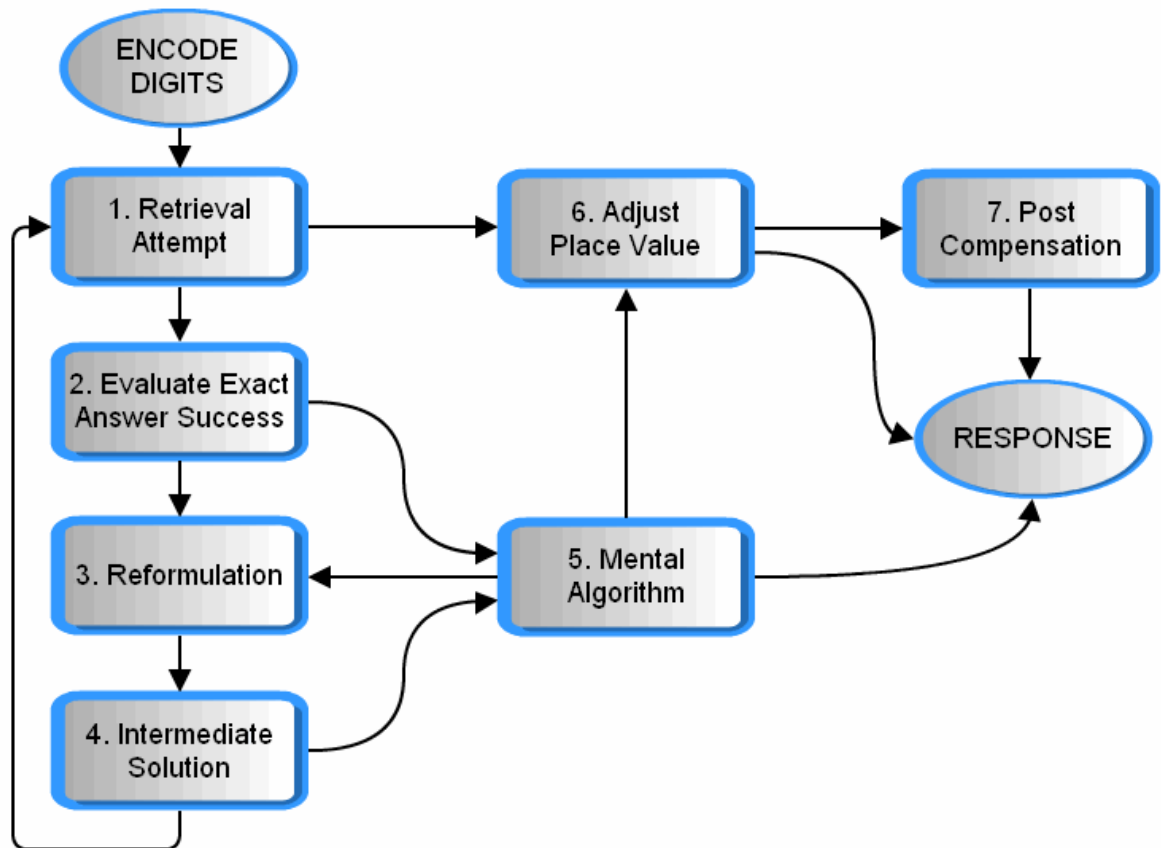


Figure 2: Process Model for Estimation (LeFevre, Greenham & Waheed, 1993)

Working Memory

Computational estimation depends upon working memory. According to prevailing theory (Baddeley, 1986; Repovš & Baddeley, 2006) all information in human memory is held for a short time in one of three working memory stores: (a) the visuospatial sketchpad, (b) the episodic buffer, and (c) the phonological loop. During

estimation, the digits of the numbers and the operations on those numbers reside in working memory, either in the episodic buffer or the visuospatial sketchpad. For digits stored phonologically using subvocal rehearsal ("saying" them repeatedly in one's mind) individuals can store up to four digits flawlessly while performing an unrelated mental task (Hitch & Baddeley, 1976). This quantity four as a limit of temporary memory storage has been confirmed by subsequent research (Cowan, 2000). Beyond seven digits (Miller, 1956) individuals are liable to make mistakes, and indeed Hitch and Baddeley (1976) found that participants misremembered over 19% of eight-digit sequences in their concurrent task study.

The ability to store more than four digits in working memory is of direct importance to an analysis of computational estimation. Miller's (1956) famous "magical number seven" paper draws a useful distinction between *bits* and *chunks* of information. A bit is the familiar *binary digit* (1 or 0) and a chunk is a unit of information as encoded in human memory, such as a decimal digit, phoneme or word. Chunks in human memory are automatically combined to form larger chunks. According to Cowan (2000) working memory is limited to four chunks, plus or minus one. Thus if more than about four digits are held in working memory while concurrent tasks are performed, it can be inferred that multiple digits have been combined into chunks.

These chunks are stored in long-term memory, from which they can be retrieved for problem solving (Baddeley, 1986). Mnemonic devices can be used to memorize extremely long strings of digits (experimentally, up to at least 79; Ericsson, Chase & Faloon, 1980). This rapid memorization of data relied upon a hierarchical chunking process, whereby the participant developed logical groups of digits and then encoded and

stored them as groups of groups, etc. Phonological encoding and rehearsing were key elements of this technique.

Strategy selection in Computational Estimation

Studies of concurrent task performance such as Hitch and Baddely's (1976) digit memorization task are particularly relevant here because computational estimation requires simultaneous manipulation of digits and strategies. Manipulation of digits becomes more cognitively demanding for certain types of problems, such as addition problems that require carrying (Kalamian & LeFevre, 2007). Working memory load has an influence on the quality of strategy selection during CE (Imbo, Duverne & Lemaire, 2006). Specifically, high load leads to selection of inferior estimation strategies. There is evidence that executive functions, specifically cognitive flexibility and inhibition ability, contribute to children's effectiveness at forming accurate computational estimates (Lemaire & Lecacheur, 2011).

Studies of computational estimation specifically (Siegler & Booth, 2004) and mathematical strategy selection more broadly among children (Siegler, 1996; Shrager & Siegler, 1998) reveal the following characteristics. Strategy selection is adaptive, so that young children will adjust strategies to match problems (Shrager & Siegler, 1998). Adoption of new strategies is gradual (Siegler & Booth, 2004) so that inefficient strategies may take years to be fully overtaken by more efficient ones. There is little reason to doubt that such phenomena as gradual strategy development and adoption persist beyond childhood; they likely explain facets of adult performance on computational estimation as well. For example, Siegler's (1996) theory of strategy

development also fits with findings of adults choosing different strategies on the same problem over multiple trials (Dowker, 1992).

Among undergraduate adults, a concurrent task will worsen performance more on computational estimation than on exact answer computation, implying that exact answer computation is more cognitively demanding (Kalamian & LeFevre, 2007). However, as described in the next section, computational estimation seems to be more cognitively demanding for children.

Developmental Changes in Computational Estimation Strategy

Individuals gain functionally greater working memory as they grow older by developing more and more efficient methods of chunking (Chi, 1978). This contributes to mathematics ability and to estimation ability. More specifically, estimation ability improves as children mature to adulthood (Case & Sowder, 1990; LeFevre et al., 1993; Lemaire & Lecacheur, 2002; Lemaire & Lecacheur, 2011; Siegler & Opfer, 2003). Very young children perceive of numbers as increasing logarithmically (Laski & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003). This makes intuitive sense because number of digits increases logarithmically with the magnitude of a whole number (integer). The result is greatly distorted estimates by young children, up to about school age.

From a neo-Piagetian perspective, the ability to estimate rather than calculate exact answers seems to appear around the formal operational stage (Case & Sowder, 1990). Case and Sowder's (1990) study with participants ranging from 6 to 16.5 years of age supports the hypothesis that estimation is more cognitively demanding than exact computation, and thus develops later. For an unpracticed estimator, there may be less

mental effort involved in operating on all original, unnecessary digits in working memory than in developing and maintaining reformulated (e.g., rounded) numbers. There was evidence that true estimation appeared around age 11, which Case and Sowder (1990) term the late *dimensional* / early *vectorial* stage, aligned with Piaget's formal operational stage.

Jerome Bruner (1966) might argue that computational estimation improves the economy of representation of the original mathematics problem by reducing the number of significant digits. Computational estimation deals strictly with symbolic representation. The use of technology to provide an iconic analog to a symbolic representation is of particular interest in this research.

Tolerance for Error

An aversion to making computational estimates persists at least to the undergraduate level (Hanson & Hogan, 2000) if not further. This manifests as a preference for exact answers and research participants' tendency to compute answers exactly even when they are asked to estimate (Sowder & Wheeler, 1989; Usiskin, 1986). This may have to do with students' low self-efficacy (Bandura, 1977) or lack of confidence in their own "nearness judgment" (Case & Sowder, 1990).

Various researchers (Dowker, 1992; Reys et al., 1982; Rubenstein, 1985; Sowder & Wheeler, 1989) suggest that a high *tolerance for error* may be necessary as an important component of estimation ability. This construct originated with Reys et al. (1982) who provided a quote from a seventh grade research participant, saying good estimators are "not bothered by being off some." Intuitively, high tolerance for error seems as though it should be a predictor of strong computational estimation ability.

However, it has yet to be operationalized in a way that correlates highly with CE ability. A study (Hogan et al., 2004) correlating personality traits with computational estimation performance found that tolerance for ambiguity does not predict strong CE performance. The same study found that agreeableness is moderately negatively correlated with CE, which Hogan et al. (2004) attributed to the skepticism associated with the negative pole of that construct.

Anchoring

Tversky and Kahneman (1974) use the term anchoring to describe the starting point for estimation. They provide evidence that anchoring influences computational estimation and percentage estimation for facts about world politics such as membership in the United Nations. The magnitude of a numerical anchor has a profound impact on a person's final estimate. For example, a percentage estimation task with anchors of 10% and 65% resulted in mean estimates of 25% and 45% for the world facts estimation task. Straight computational estimation tasks were even more heavily influenced by the order of the operands. For example, the computational estimation problem $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$ resulted in a mean estimate of 512, while the problem $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ resulted in a mean estimate of 2,250. (The exact solution to this problem is 40,320.) The first digit seems to anchor the remaining mental computation to a significant degree.

Computational Estimation in Education Literature

In their seminal study on the processes used for computational estimation, Reys et al. (1980) classified several processes used by skilled estimators. They named the three key processes reformulation, translation and compensation. These processes could be used separately or together, and could be repeated up to several times each during the

solution of a single computational estimation problem. A process is defined here, following Reys et al. (1980) as a category encompassing several strategies. Several strategies, which are more specific fine-grained mechanisms for estimation, are defined within each process below.

Process: Reformulation

Reformulation consists of altering the numbers involved in the problem, and can involve rounding and truncation, among other strategies. Rounding generally depends upon the digit that is to be the greatest non-significant digit. For example, if a three-digit number such as 352 is to be rounded so as to have only one significant digit, it can be rounded up to 400 or down to 300. Very frequently, rounding is taught so that the 5 in 352 is used to decide whether to round up or down. Digits 5 and greater signal an estimator to round up and digits 4 and less signal rounding down. Hereafter, this strategy is referred to as algorithmic rounding. It is commonplace in U.S. classrooms (Reys & Bestgen, 1981). If the rounded value is to have two significant digits (e.g., 350) then the third digit is used, following the same criterion. Reys et al. (1982) refer to any strategy that depends on most-significant digits as a front end digits strategy.

A slightly simpler front end digits strategy, which they refer to as truncation, is the same as rounding down, ignoring the value of the greatest non-significant digit. For example, truncation to the front end digit of any integer in the range 301 to 399 results in 300. In general, rounding or truncating to a multiple of 10 simplifies the problem by reducing the number of significant digits, but rounding can also be done to some other value with which the estimator has fluency. For example, an estimator may round to the nearest multiple of 5.

Reys et al. (1982) distinguish between operations on just the most significant digits, while ignoring the trailing zeroes that resulted from rounding or truncation, and operating with all of the digits. Operating with all of the digits is referred to as Round-SND or Truncate-SND (for Same Number of Digits). Operating with just the most significant digits is referred to as Round-EXT or Truncate-EXT (for EXTracted digits).

Reys et al. (1982) also use the term "compatible numbers," meaning numbers that sum to a multiple of 10 or that are multiples of each other, to simplify division. Estimators can reformulate fractions to decimals or vice versa, and as mentioned above can operate with multiples of numbers with which they feel comfortable.

Process: Translation

Translation involves altering the "mathematical structure" of the problem, such as changing repeated addition of similar numbers to multiplication. For example, a person forming an estimate for $8,946 + 7,212 + 7,814$ might recognize that the three numbers are all roughly centered on 8,000, and thus provide $3 \times 8,000 = 24,000$ as an estimate. This kind of strategy generally involves changing the operation being performed on the numbers as well as changing the numbers themselves.

Dowker (1992) reports several highly sophisticated examples of translation demonstrated by professional mathematicians performing an estimation exercise. One mathematician, when asked to estimate 12.6×11.4 , used the $(a + b)(a - b)$ rule. Others transformed numbers into powers of other numbers such as translating 64.6×0.16 into $2^6 \times 2^4 = 2^{10}$ and then adjusting place value.

Process: Compensation

Reys et al. (1982) also noted that estimators sometimes adjust values mid-course or afterward. They named these processes intermediate- and final-compensation, respectively. Intermediate compensation involves recognizing that an estimate will be too large or too small even before performing the operation. For example, if two numbers are both to be rounded up the estimate is likely to be unreasonably large. Thus as an intermediate compensation an estimator may round one of the numbers down. In another example that Reys et al. (1982) provide, a 9th grader is asked to estimate the sum of six numbers. The student rounds four of the numbers up, and compensates by eliminating one of the numbers. Final-compensation involves a similar technique, applied after the operation has been performed. For example, estimating $21,319,908 \div 26$ as $26,000,000 \div 26 = 1,000,000$ is an overestimate, which can be compensated down to 850,000 (Reys et al., 1982, p. 190).

Levine's Strategy Classification Scheme

In the same year that Reys et al. (1982) published their paper describing processes used by good estimators, Deborah Levine published a paper that classified estimation strategies (1982). Levine (1982) describes eight categories, resulting in a more precise but less parsimonious classification scheme than Reys et al. (1982). All of the strategies she identifies are described in the above sections except for the strategy of *proceeding algorithmically*. This means following traditional methods for long division or multi-digit multiplication using reformulated numbers. It is sometimes useful to know when such a rule-based strategy with intermediate steps is being used. Levine's (1982) scheme has been used in several subsequent studies (e.g., Dowker, Flood, Griffiths, Harriss &

Hook, 1996). However for parsimony, a modified version of the Reys et al. (1982) scheme is used in this research.

Teaching Computational Estimation

Despite the breadth of the literature on estimation, there are few empirical studies that address whether and how estimation abilities can be improved. Fung and Latulippe (2010) describe a promising study in which pre-service teachers showed significant learning gains on a computational estimation task, but the authors do not provide information on the duration or exact nature of the intervention. Another recent study (Lan et al., 2010) showed significant gains among fourth graders after a four-week series of lessons on estimation.

Lan et al. (2010) implemented Reys' (1986) recommended strategy for teaching estimation. While Reys' (1986) article makes useful distinctions and describes techniques for teaching estimation, it is not an empirical study.

Motivating Computational Estimation in the Classroom

One of the reported difficulties of teaching computational estimation is that students in a classroom will often calculate an exact answer when they are asked to form an estimate (Usiskin, 1986). Usiskin (1986) reports that students will round an exact answer to reduce its significant digits and provide the rounded result as an estimate. Other mathematics education researchers confirm the use of this strategy (Sowder & Wheeler, 1989). For example, when asked to estimate 49×82 students will find the exact answer of 4018, and then report 4018 or perhaps 4000 as an estimate. For the purposes of this research an exact answer, such as 4018 in this example, will be considered to be an estimate (with 0% error) as long as it is calculated mentally. This is a

practical strategy because in real world situations, an exact answer is generally just as useful an approximation, or more so. Possible exceptions, as Usiskin (1986) points out, occur when additional significant digits are unnecessary or unwarranted. Computational estimation is intended for simplification, but it involves greater abstraction than straight mental computation (Case & Sowder, 1990) and thus the cognitive benefits of estimation can be lost on students. Rounding an exact answer to something inexact (such as turning 4018 into 4000) sacrifices both simplicity and precision and so provides neither benefit. This technique will be referred to as *pseudo-estimation*.

To limit the use of pseudo-estimation and to provide a broad context for the need for computational estimation, it is important for K-12 educators to motivate the use of computational estimation. This can be done by providing examples of the real world need for estimates. Often, the constraint of time can also be used to encourage students to make estimates.

Students' Difficulties with Computational Estimation

To a novice, computational estimation can be more complex than straight computation (Lefevre et al., 1993) because of the extra processes involved (Case & Sowder, 1990). Malcolm and Berry (2011) frame the problem of teaching estimation as a paradox because making a reasonable estimate depends upon having a sense of the magnitude of the exact answer, which in turn depends on making a reasonable estimate.

Declarative Components of Computational Estimation Ability

The first and most obvious skill to teach children computational estimation is a working definition of the term estimation. Students should be able to explain what estimation is, as well as when and why it should be used. The closely related idea of

reasonableness is one that is important and often missing from the curriculum at many levels of current K-12 education (Rheinlander et al., 2008).

LeFevre et al. (1993) analyze the knowledge necessary for computational estimation as procedural and conceptual components. These provide useful KCs for teaching computational estimation. Their first two guiding principles for estimation are *simplification* and *proximity*. Students should understand that the primary purpose of estimation is to simplify problem solving. This will require a tradeoff with the principle of proximity, because estimates that are closer to the exact answer are generally more difficult to compute. For example, the more significant digits that a person retains while making an estimate the more complicated the mental processes of estimation will be.

Procedural Components of Computational Estimation Ability

Estimation skill is highly correlated with and dependent upon other math skills, particularly the ability to work with place value (Rubenstein, 1985). Emphasizing this in the mathematics classroom is certainly important. Moreover, teaching certain targeted strategies for estimation may help. For example, a round one up and one down strategy, also known as a mixed-rounding (Lemaire & Lecacheur, 2011) or initial compensation (Reys et al., 1982) can often provide better estimates than algorithmic rounding based on the second significant digit.

The power law of practice (Newell & Rosenbloom, 1980) supports the common-sense notion that performance improves with practice. A goal of the current research is to investigate whether particular components of estimation contribute significantly to improvement.

Sources of Error in Computational Estimation

Estimation by its nature implies that there will be some error. Good estimation strikes a balance between simplicity of mental calculation with low error. Large errors can result from a failure to encode the problem correctly, such as misreading the digits or the operators. Large errors can also result from a faulty choice of strategy (Threadgill-Sowder, 1984). Finally, large errors can result from a faulty execution of strategy, either through a mathematical or memory error. This last source of errors is usually a result of working memory load. Anderson (1993; p. 240) points out that when learners make errors, these can be explained in the following three ways:

1. Learners know correct production rules, but apply them incorrectly.
2. Learners do not know correct production rules, and therefore guess.
3. Learners know incorrect production rules.

This is a straightforward way to explain errors. Butler and Winne (1995) describe a five-part model of feedback that is designed to address these same three possibilities. Estimation has a range of reasonable answers whose width varies depending upon context. Thus, the corresponding three ways learners would make unreasonably large errors in estimation would be:

1. Learners know good computational estimation strategies but make mathematical errors.
2. Learners do not know relevant computational estimation strategies and therefore guess wildly.
3. Learners have a faulty understanding of computational estimation and consistently select strategies that lead to unreasonable estimates.

The first type of error here could be mapped on to any one of the errors described above by Anderson's (1993) classification of errors, as it applies to mathematics. In general, Anderson's first type of error ("Learners know correct production rules, but apply them incorrectly") can be further broken down into what Anderson (1993) calls slips such as typos, or incorrect application of finer-grained production rules or declarative knowledge. To take a concrete example, a student might solve 18×19 by rounding 19 to 20 and then multiplying 18×20 . This strategy usually results in an estimate that is reasonable for most real world applications. However, if a student thinks that 28 is the product of 18×2 , then he or she will provide 280 as an estimate. This is an example where the student has a good grasp of the procedural steps (i.e., KCs or production rules) of 2×2 multiplication estimation in general but has faulty declarative knowledge of the product 18×2 .

Context and Reasonableness in Estimation

Context plays a particularly important role in estimation, where the acceptable range of a reasonable answer depends upon the problem itself. Garofalo and Bryant (1992) describe a true example of a help-desk telephone company representative responding to a \$60,000 overcharge on a phone bill by asking the unfortunate recipient of the bill to pay the full amount, and to be assured that "any excess payment would be credited toward her next bill". In this case, reflecting on the fact that a phone bill is usually close to \$60 would have helped the help-desk representative understand the unreasonableness of a \$60,000 bill.

This is an extreme example, and there are few real-world situations in which a 99,900% error would go unnoticed. However, the precision of an estimate depends

sensitively on the situation (Usiskin, 1986). A customer's mental estimate for the cost of a hamburger may reasonably be much less precise than a newscaster's reported estimate for a record-breaking runner's finish time in terms of significant digits.

LeFevre et al. (1993) found that some middle-school aged students actually made better computational estimates for word problems than for "naked number" (no-context) problems. They postulated that the context provided by the word problems enhanced these students' abilities to make reasonable estimates. This trend was very slight, non-significant and not consistent across grade levels, but suggestive of interacting components within estimation problems.

Context is crucial to estimation, so digital technology interfaces must be carefully designed so as not to strip context from computational estimation. The next sections deal with a critical component of technology for learning: feedback. This is one of the central manipulations in the present research.

Feedback in Digital Technology

Digital technology-based feedback as it is used in education is a double-edged sword. Technology can be either a valuable resource or, if it is overused, a crutch. A meta-analysis of calculator use in K-12 classrooms (Ellington, 2003) demonstrated that calculator use does not hinder the development of mathematics abilities. Disturbingly, however, research by Berry (1998) showed that students will unquestioningly accept output on a calculator, even when output is modified to differ from an exact answer by as much as 50%. A broader look at the definition and research on feedback itself will inform this research.

Feedback is a very broad topic that has been defined and studied in a variety of ways (Butler & Winne, 1995; Corbett & Anderson, 2001; Hattie & Timperley, 2007). The following sections establish the broadest working definition of feedback that is available in the literature. Next the definition is refined to pertain to learning via computer and then to mathematics learning in particular. Feedback can vary along various spectra. The two dimensions that seem most relevant to the research here are immediate vs. delayed and complex vs. simple. The following analysis addresses the point along each spectrum that would logically provide the most benefit to learning estimation abilities.

These sections draw heavily upon research with cognitive tutors (Anderson, 1993; Corbett & Anderson 1991), which have proven to be fertile ground for studying proper feedback conditions in teaching computer programming and mathematics. The earliest cognitive tutors were based on core principles that included providing immediate feedback (Anderson et al., 1995). Subsequent research has suggested that a more nuanced approach to the timing of feedback may be more effective for learning.

Definition of Feedback

Feedback provides a means to update or modify knowledge. Butler and Winne (1995) define feedback as "information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory." They specify "information in memory" following Alexander, Schallert and Hare (1991) as domain knowledge or metacognitive knowledge. The five functions they list correspond to five learner states. Confirming feedback is provided in response to actions demonstrating accurate knowledge. Add-to, overwrite, tune and restructure types of feedback are responses to

incomplete, inaccurate, somewhat inaccurate and poorly structured knowledge, respectively.

Proper Timing of Feedback

One key question in the design of feedback in an online learning system is when the feedback should be provided. Schmidt, Young, Swinnen and Shapiro (1989) posit that immediate feedback benefits the efficiency of problem solving while delayed feedback improves deep learning, as demonstrated by later retention and the ability to transfer learned knowledge.

For teaching complex problem solving activities, ongoing feedback benefits learners more than feedback that is withheld until the end of each problem (Corbett & Anderson, 2001). Corbalan, Paas and Cuypers (2010) found that learners solving multistep mathematics problems online prefer computer-generated feedback at each step in a problem over receiving feedback just at the end. In general, Schmidt et al.'s (1989) argument that immediate feedback improves the efficiency of problem solving is supported elsewhere (Corbett and Anderson, 2001). However this efficiency may come at the cost of depth of student understanding of the material.

Nathan (1998) states that learners who receive too much feedback too quickly can come to hold an oversimplified understanding of the problem solving process. To avoid this oversimplified understanding, Hattie and Timperley (2007) suggest that feedback is most effective when it accounts for the interaction between timing and complexity. That is, simple task-related feedback should be provided immediately and complex process-related feedback should be delayed. A correlate of this idea is the idea that dramatic differences arise in learning from a reading task when delay is held constant and problem

complexity is changed (Clariana, Wagner, & Roher Murphy, 2000). Learners improve with delayed feedback on complex problems and actually do worse on simple problems with the same feedback delay. This leads to the following discussion of the complexity and detail of feedback.

Proper Complexity and Detail of Feedback

The literature on the relationship between effectiveness of feedback and its level of complexity and detail paints a nuanced picture. Hattie and Timperley (2007) draw a distinction between task-based feedback (FT) and process-based feedback (FP). Task-based feedback is focused on specific details of a learner's current state in problem solving. Process-based feedback looks more broadly at a learner's approach to solving a class of problems. By its nature, FP must be more abstract than FT, and thus more conceptually complex. One benefit Hattie and Timperley (2007) cite for the use of simpler, task-based feedback (FT) is that it can enhance confidence, which can in turn result in spontaneous improvements in deeper understanding (Earley, Northcraft, Lee, & Lituchy, 1990).

Some research supports this conclusion that feedback should be simple (Mathan, 2003). Mathan's (2003) developed an *intelligent novice* model, an adaptation to the established pattern underlying cognitive tutors (Anderson, 1993) which Mathan referred to as the *expert* model by contrast. The intelligent novice model depends upon detailed, delayed feedback that let students make some errors and then reflect on those. The first implementation of this model probably incorporated too much detailed, complex feedback and resulted in learning that was not appreciably better than the traditional

expert model. However, a revised version of intelligent novice model showed significant gains both on pre-post

Other research indicates that feedback should be highly detailed (e.g., McKendree, 1990). Corbalan, Paas and Cuypers (2010) found some evidence to suggest that learners prefer more feedback on more difficult problems. This compliments research (Halabi, 2006) showing that novices benefit more from rich and detailed feedback in similar environments.

Conclusion: Interaction with Learner Knowledge

In sum, a pattern emerges, where more novice learners need more immediate and more detailed feedback, while more expert learners generally need feedback to be brief and infrequent. This paints a picture in which learners in a new domain need large amounts of feedback quickly to establish basic understandings, and experts need only the occasional feedback to "stay on track". There may well be a middle ground between novice and expert – a journeyman level – during which delayed feedback fosters the development of evaluative skills (Schmidt et al., 1989). These evaluative skills are largely metacognitive and are necessary for independent problem solving.

The knowledge-feedback interaction has direct implications for instructional design. An important concept that is operationalized in cognitive tutors is the concept of floundering. Anderson et al. (1995) define floundering as either a) three errors in a row or b) two consecutive errors that the tutor cannot explain internally. The tutors described by Anderson et al. (1995) gave help either when students began to flounder or when they asked for help (p. 9). Novices are apt to flounder if they are provided with insufficient feedback.

Gaming the System with Feedback

If learners are allowed to simply probe the instructional system for the complete feedback on each problem, they can receive too much help. Baker et al. (2009) describe what they call gaming the system, i.e. depending upon built-in help in cognitive tutors as an easy source of correct answers. This represents a misuse of the feedback in cognitive tutors that is designed to be used occasionally or as a last-resort.

Giving versus Withholding Assistance

Koedinger and Aleven (2007) respond to the issue of students gaming the help system of cognitive tutors by means of the mastery learning system. In cognitive tutoring terms, asking for help until the answer is provided is called bottoming-out, using the analogy that students work through hints from top to bottom. The mastery learning system monitors student progress and gives students problems targeted to their level of mastery. Students who game the help system will not progress through mastery levels. Instead they will be given repeated problems of the same type each time they bottom out the help system.

Conclusion

The present research concerns middle school students, who should have a moderate understanding of computational estimation. They might not know the processes described here, but will be developmentally capable of learning them (Case & Sowder, 1990). They may well have a good intuition for the tradeoffs between simplicity and proximity (LeFevre et al., 1993) that computational estimation entails. The use of anchors (Tversky & Kahneman, 1974) for a given problem context will help students in developing a sense of what is a reasonable estimate in that context. This intuition for

what makes a reasonable estimate likely contributes to previously unseen contributors to estimation ability (Rubenstein, 1985).

The technology that is used in this research will provide instant feedback for students on their computational and measurement estimates. This may be subject to gaming behavior (Baker et al., 2009). The next chapter introduces the Estimation Calculator and the Spatial Estimator, two technology tools that have been used in studies leading up to the present research. The form of feedback in these tools is instantaneous and simple. The feedback has been used as a catalyst for discussion between students and teachers on how students approach computational estimation. Until now, its effects on learning have not been studied.

Chapter 3

The Estimation Calculator

The Estimation Calculator is a variation on a four-function calculator designed for practicing computational estimation. It currently exists in a web-based format and as an iOS app for the iPod Touch, iPad and iPhone as iEstimation. The calculator provides feedback in the form of a bi-directional error bar showing the relative error of a user's estimate, and provides the exact answer to a problem only when a user's estimate is within a given tolerance. By default, this tolerance is set to $\pm 15\%$ of the exact answer. As a visual indicator of this tolerance, there is a red box around the center of the bi-directional error bar whose width represents its proportion of the error bar (see Figure 3).

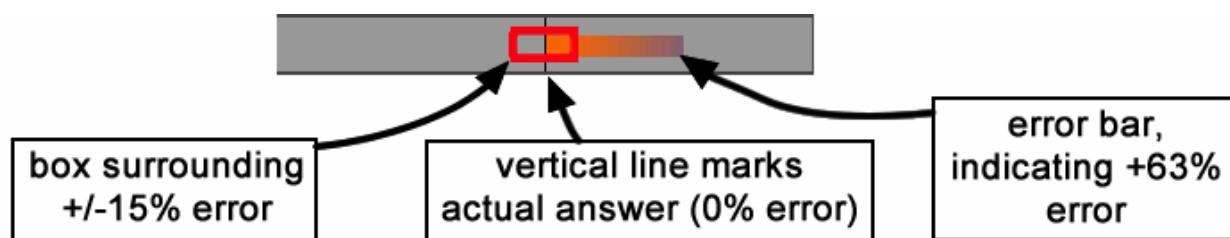


Figure 3: An Illustration of the Error Bar on the Estimation Calculator

The Estimation Calculator was developed as a method to strengthen students' understanding of the reasonableness of an answer (Garofalo & Bryant, 1992; Rheinlander et al., 2008). Computational estimation requires thinking at a level of abstraction (or dimension; Case & Sowder, 1990) that is beyond standard computation of arithmetic problems and is developmentally accessible for upper elementary and middle school aged students. Students in this age range, roughly between ages 8 and 13, are the target learning population for the Estimation Calculator. The Estimation Calculator provides

automatic feedback on an estimate, which allows a student to compare his or her estimate, which persists on the display, to its relative error. A student can then synthesize the estimate, the original problem and the relative error to revise the estimate. The student may do this iteratively to converge upon a reasonable estimate.

The following two sections describe studies conducted on the Estimation Calculator to explore its use for students to practice computational estimation.

Study 1: Exploratory Research with the Estimation Calculator

Objective

The first study was a pilot to answer the following question: How do students and tutors use the Estimation Calculator in practice? The initial hypothesized model was the minimal description of how students might use the Estimation Calculator (See Figure 4).

This four-step model included:

1. simplify problem
2. solve simplified problem
3. check estimate on Estimation Calculator
4. revise estimate

It was predicted that students would iterate on the third and fourth steps in this model to revise estimates as necessary. The motivation for the research was to expand and elaborate on this model.

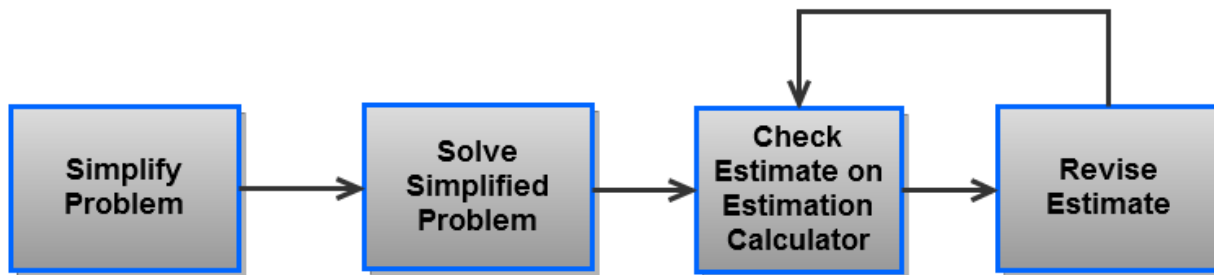


Figure 4: Initial Process Model for Estimation Calculator

Participants

Five pre-service elementary teachers and nine fourth grade students from a public elementary school in central Virginia participated in this exploratory study with the Estimation Calculator. Pre-service teachers were intended to have two half-hour online meetings with students to tutor them in whole number computational estimation strategies. However, four new students were brought into the study for the second session, replacing four of the students from the first session. This was done to broaden the participation and give more students the opportunity to receive tutoring.

Research Design

This study was designed in an interview format to provide qualitative data. Pre-service teachers in a combined Bachelors and Masters of Teaching program met with fourth grade students in a virtual environment. There the Estimation Calculator version for the web was screen-shared between the pre-service teachers and the students. Pre-service teachers received a series of problems on paper along with an interview protocol that had been developed by the investigators. Pre-service teachers were encouraged to ask probing questions and even to suggest estimation approaches but were discouraged from prompting students with actual estimates (such as "why not try 14,000").

Online Interface

The pre-service teachers met with the students online through Blackboard Collaborate for two half-hour sessions. During these online sessions, the pre-service teachers and students were physically separated but could also see one another via web-cameras. They screen-shared the Estimation Calculator web version through Collaborate, with the pre-service teachers controlling the input such as typing in estimates and the students providing spoken instructions (See Figure 5). The intention behind this design was to elicit student thinking as they solved the problems.

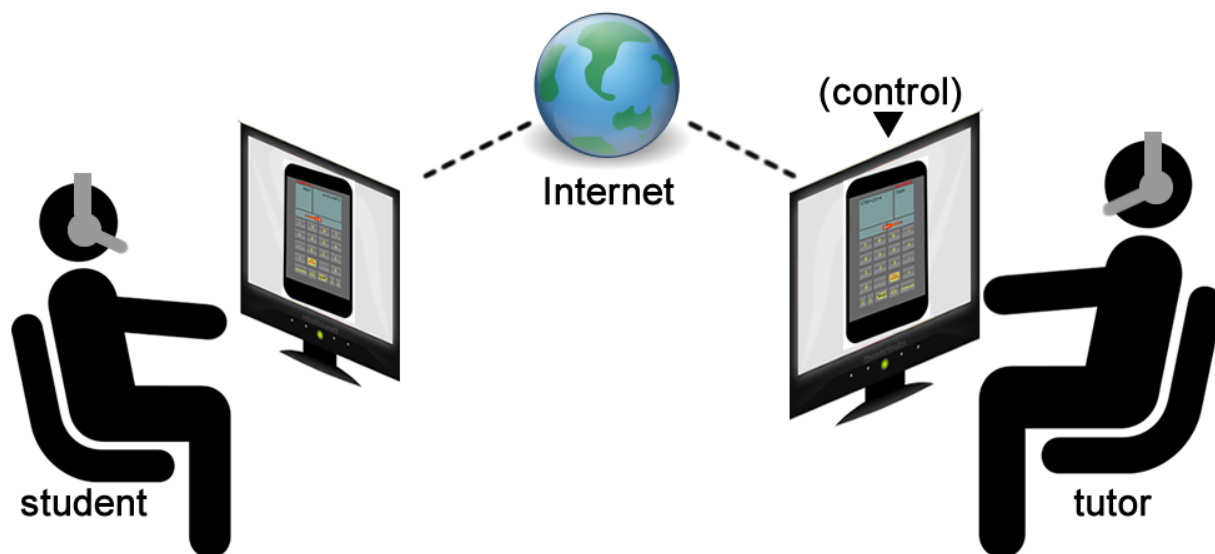


Figure 5: Remote tutoring link scheme

Problem Set and Interview Protocol

The questions in the problem set included items derived from the Assessing Computational Estimation (ACE) test (Reys et al., 1982) as well as new problems. All were whole number computational estimation problems using the four basic operations:

addition, subtraction, multiplication and division. The first seven problems were addition and subtraction problems with operands between two and five digits in length. Some problems included two large operands, such as $87,419 + 92,765$. Others included three small operands, such as $127 + 382 + 81$. In each case, the problems were designed to exceed the four to seven digit working memory capacity (Cowan, 2000; Miller, 1956) and to elicit a variety of estimation strategies. The full problem set is listed as Appendix A. The interview protocol included asking for students to verbalize their strategies. It also included questions about how confident students felt about their estimates before and after receiving estimate error feedback (EEF) shown on the error bar. The pre-service teachers asked students whether the estimate was above or below the exact answer, and had students revise estimate accordingly.

Analysis

Screencasts with audio were recorded from the interviews. Of the ten original interviews, three were unavailable because of technical difficulties. The remaining seven were transcribed and analyzed to establish whether the model described above (see Figure 4) adequately described the students' behavior.

Results

Four distinct patterns of behavior emerged. The process model was updated to incorporate these behaviors. The resulting, updated process model is shown in Figure 6. The behavior patterns included (1) masterful estimation on the first attempt, (2) estimate

revision after a few attempts, (3) floundering¹, and (4) strategy revision after many attempts. The following dialogue samples, with pseudonyms, provide examples of the three behavior patterns with the Estimation Calculator.

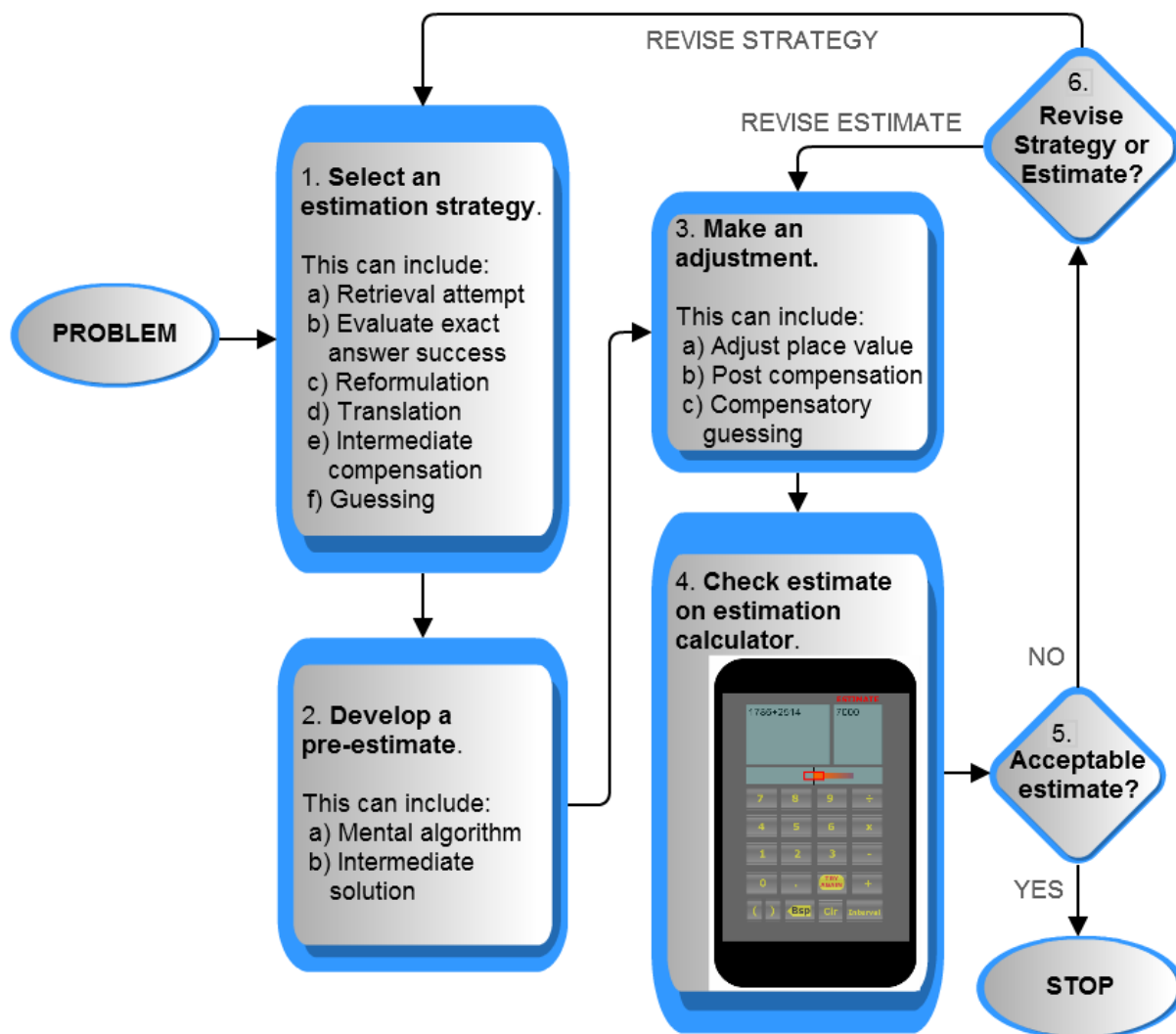


Figure 6: Revised Process Model

¹ Floundering was not included as a behavior pattern in the original publication on this study (Malcolm & Berry, 2011) although there was evidence of it.

Masterful First Attempt at Estimation

The first dialogue sample between Rick, the pre-service teacher, and Cindy, the student, demonstrates masterful estimation on the first attempt:

For the problem $367+748$:

Rick: "So, what do you think on this one?"

Cindy: "Um well, 748 is bigger than 367, so I'm thinking ... that the estimate will be close, around 1000."

Rick: "Good. And why do you think that?"

Cindy: "Because if you add the seven and the three together then that would equal one thousand."

Rick: "Good."

Cindy: "And if you add the rest of it, then it would be more than one thousand, so it would be closer to one-thousand..."

Rick: "Good. So should I put in one thousand?"

Cindy: "Yes."

Rick: Good. Nice. So first try you got it. You got it within the range they want.

So what's the actual answer?"

Cindy: "One thousand, one hundred fifteen."

This interaction demonstrates the successful use of a truncation and front-end digit strategy. Cindy seems to have rounded both operands down to the nearest multiple of 100. This technique of using only left-most digits without adjustment for less significant digits is considered to be truncation (Reys et al., 1982). Because she describes her solution as "adding the seven and the three" it can be inferred that she performed the

calculation with these front-end digits by themselves and then mentally adjusted place value afterward for her estimate of 1000.

Estimate Revision after One or More Attempts

For this first kind of behavior pattern, the Estimation Calculator reinforces for the student that he or she has performed the estimation successfully. However, the second behavior pattern – revision of estimates – was equally common among the students in the study. The participating group of students had difficulties working with place value and multiples of ten. This was the cause of error for many students' estimates, as is demonstrated in the following conversation. Here the tutor, Gerry, is working with student, Zara, to develop an estimate for the problem $42,163 \div 73$.

Zara: "That would be forty-thousand divided by seventy."

Gerry: "Ok."

Zara: "So that would be ... around five, I think."

Gerry: "Ok. So, five ... what?"

Zara: "Maybe, fifty?"

Gerry: "All right. Let's try it."

(The error bar appears at -91%)

Zara: "Maybe we should add another zero."

Gerry: "So, you want to add one zero? So that would be five-hundred?"

(Zara nods. Gerry keys it in. The correct answer appears.)

Gerry: "Oh. Look!"

(The error bar appears at -13%, within the acceptable tolerance, and the correct answer of 577.57 appears on the calculator.)

This dialogue provides an excellent example of what Reys et al. (1982) refer to as compensation. In this case, the feedback on the Estimation Calculator alerts Zara to the need to revise her estimate. Zara senses that her strategy is valid, although something has gone wrong. She quickly decides compensate by adding zeroes and converges on a reasonable estimate. The resulting estimate had a reasonable most-significant-digit (five) and a reasonable magnitude (on the order of 100).

This dialogue also highlights one of the optimal uses of the Estimation Calculator. The difficulty of estimation as an abstraction (Case & Sowder, 1990) and as a load on working memory (Lefevre et al., 1993) can be alleviated by this incremental behavior pattern. Zara solves this problem one piece at a time, manipulating the most significant digits first. Her original estimate of 5 is the number of times that 7 can be divided into 40, ignoring the remainder. This may also have been a random guess, but other conversations with Zara suggest that she is capable of making this kind of calculation quickly and accurately. At this point, her working memory seems to be taxed. After a slight adjustment, to 50, she asks Gerri to enter the problem into the Estimation Calculator. The information displayed gives her a cue to increase her estimate.

Zara had difficulties with place value on other problems as well. On multiple occasions she made estimates with reasonable front end digits but unreasonable magnitude. The Estimation Calculator allowed her to tackle these difficulties as a distinct procedural step rather than concurrently. To put this in cognitive load theory terms, the Estimation Calculator reduced the intrinsic cognitive load of these problems. This step-wise approach can also be seen as a conversion of the germane cognitive load of magnitude (number of digits) into extraneous cognitive load. Students can ignore

magnitude while they manipulate front-end digits, as Zara does in the example above. The Estimation Calculator feedback will alert students to errors of magnitude, so that students can then correct this error separately.

Floundering

Anderson et al. (1995) define floundering with a Cognitive Tutor as repeating the same kind of mistake three times or making two mistakes that the tutor does not recognize (p. 174). A comparable situation occurs with the Estimation Calculator when a student makes three or more unreasonable estimation attempts. This is classified here as a distinct behavior pattern, and occurred on several occasions in this study despite (or occasionally, because of) the tutoring from the pre-service teacher. In the following example, Gerri, the pre-service teacher from the previous dialogue sample, tutors a student Melissa to estimate a solution to 62×23 :

Gerri: "Sixty-two times twenty-three."

(The student, Melissa, hesitates.)

Gerri: "What do we want to round the 62 to?"

Melissa: "Sixty."

Gerri: "Good."

Melissa: "... and ...23 to twenty."

Gerri: "Ok. So, sixty times twenty. What do we think that is?"

Melissa: (thinks) "Five. Well, fifty."

Gerri: "Ok. Let's try fifty." (She types.) "Oops. Way too low. So what are we going to do when it's too low?"

Melissa: "Add another zero."

...

The dialogue continues in a similar fashion. Melissa and Gerri make a total of seven estimates before converging on an estimate that is reasonable. The sequence of estimates is: 50; 500; 5,000; 1,000; 2,000; 1,900; and finally 1,400. Melissa makes an initial estimate of five, which she quickly corrects to 50, just as in the previous dialogue between Gerri and Zara. In fact, Melissa is probably anchoring (Kahneman & Tversky, 1974) on five because it was the most significant digit for a reasonable estimate to the previous problem. As a strategy, "choose the same digits as in the previous answer" is not likely to be more successful than random guessing. Melissa seems to be making this guess because she is overwhelmed by trying to multiply 60×20 . Several front-end strategies (Reys et al., 1982) such as 6×2 , or 6×20 , or 60×2 , would likely have resulted in a lower cognitive load for Melissa. Lower cognitive load might have resulted in faster success.

Strategy Revision After Repeated Attempts

The floundering behavior pattern represents one of the pitfalls of the Estimation Calculator. After making an initial unreasonable estimate, students can use a compensation strategy to adjust the magnitude, as Melissa and Gerri did for the first three estimates (50; 500; 5,000). None of these three estimates is based on a well-reasoned front-end strategy, and Melissa and Gerri "overshoot" the tolerance region by appending zeroes. Their focus on the output of the Estimation Calculator drives the rest of the discussion. The remaining four estimates are all based on correcting the error shown on the display. This behavior is categorized as floundering because it is inefficient and does not focus Melissa's attention on the important parts of the problem.

There is no dialogue snippet for the strategy revision behavior pattern because there was no evidence of it in this study. This behavior pattern is provided as suggested "best practice" for a tutor working with a student. If a student makes three or more unreasonable estimates, it is recommended that a tutor should intervene, and refocus the student's attention on the numbers in the original problem. In the above scenario, Gerri could have reintroduced Melissa's original reformulation of the problem (60×20) and elicited new strategies. That would have been an example of strategy revision, a process that the student could eventually internalize in order to improve at computational estimation.

Discussion

Based on the results of this study, a new process model was developed (See Figure 6). This model includes the following steps:

PROBLEM

1. Select an estimation strategy
2. Develop a pre-estimate
3. Make an adjustment
4. Check estimate on estimation calculator
5. Acceptable estimate?
6. Revise strategy or estimate?

STOP

This model incorporates adjustment and post-compensation as in the Lefevre et al. (1993) model as well as the processes described by Reys et al. (1980). Both steps 5 and 6 in this model are decision steps. If the first estimate is reasonable (step 5) then no step is

visited more than once. This is masterful first attempt behavior pattern described above. If a student's first estimate is not reasonable (step 5), then he or she must make a choice in step 6. The student may revise the estimate itself with the help of the error bar (step 3) or revisit the original numbers in the problem and revise the strategy (step 1). Overuse of the error bar can result in floundering, i.e. looping through steps 3, 4, 5 & 6 repeatedly. After three of these loops it is suggested that a tutor intervene to encourage a revision of the strategy. This is a return to step 1 in the model and a revision of the strategy based on the original numbers. In the description above, had Gerry reintroduced Melissa's reformulation of 62×23 to 60×20 after the floundering estimates of 50, 500 and 5,000, that would have been a return to step 1 in the model. What followed would have been strategy revision.

Study 2: Refining the Estimation Process Model

Objectives

This study explored the question: How do students use the Estimation Calculator in practice? It was conducted with the aim of (1) devising a reliable set of problems for use of the Estimation Calculator and (2) establishing a robust coding scheme to describe the process of estimation with the Estimation Calculator. There were two parts to the study, a timed estimation component and a one-on-one Estimation Calculator activity. Whereas in Study 1 the Estimation Calculator activity with students was treated as a tutoring session, in Study 2 it was designed to be strictly an interview. The interview protocol did not include providing hints or strategic suggestions to students.

Participants

Participants were six fourth and fifth grade students at a private school in central Virginia. At this small school the fourth and fifth grade students were taught mathematics together in a combined classroom. Students had recently received lessons on computational estimation in this class. The combined fourth and fifth grades included more than six students, but the sample size was limited by the number of students who agreed to participate and returned institutional review board (IRB) release forms. All six students were age 10 at the time of the interviews.

Research Design

The first part of the research in study 2 was a timed estimation exercise. Students were given the series of problems in a timed PowerPoint projected in the front of a classroom, with 15 seconds to answer each of the problems. These problems did not include the practice problem, so the estimation with the PowerPoint began with problem #1. Questions #4, #8 and #12 are recovery items (Olshen, 1975, as cited in Reys et al., 1980). Recovery items are designed to be easier than other items, to allow students to recover cognitively from challenging problems. The standard (non-recovery) items were strictly addition and subtraction problems of five-digit numbers generated at random.

Two researchers visited the school and administered the timed PowerPoint as part of the students' regular mathematics class. After the timed PowerPoint, the researchers discussed the Estimation Calculator with the students. Researchers returned two weeks later to conduct the interview portion of the study.

The interviewers were graduate students with backgrounds in mathematics education and computer science, respectively. Both had familiarity with the functioning

of the Estimation Calculator and with estimation strategies. Each interviewer worked through the same series of problems, including the practice problem at the beginning. This problem was designed as a relatively simple introduction to the functioning of the Estimation Calculator. Interviews, which lasted between 10 and 15 minutes, were conducted individually in quiet room in the school as pull-outs from the students' regular class activity. Audio recordings and screencasts of these interviews were captured using Camtasia Studio. As in Study 1, students advised the interviewers on which numbers to enter into the Estimation Calculator and when to click the check estimate button.

Analysis

Although written estimates from the timed PowerPoint estimation component of the study were collected, these were largely incomplete. Altogether, the six students in the study only completed 51% of the questions. There was a general feeling of apprehension at the speededness of the questions. One of the students (not one of the research participants) broke down in tears. Because of the incomplete nature of this data and the apparent stress caused by the activity, this data was not included in further analysis. Future studies on computational estimation with this age group will include more explicit briefing and segments of more than 15 seconds to complete estimation questions of this complexity.

The problems used in the interview component two weeks later were the same problems from the timed PowerPoint, with the addition of the practice problem. The ten to fifteen minute interviews allowed for discussion of four to six problems, including the practice problem. Analysis of these problems was conducted using the process model developed in Study 1 (Figure 6). Two researchers conducted the analysis. One of these

researchers had also conducted interviews. The other has a background as an elementary teacher and instructional technology resource teacher. The two researchers coded one of the interviews together and consulted one another for questions on a second interview.

Results

A highly simplified version of an item analysis (Allen & Yen, 1979) was performed on the first four items from the interview (see Table 1). These were the items that were completed by all six students. Difficulty was calculated as number of attempts students made before arriving at a reasonable estimate, unlike the traditional difficulty measure of percent correct across a group. The item analysis did indicate that item 1 should be removed from the problem set. Students tended to round one of the two operands in that problem (98,073) up to 100,000, resulting in an extra digit. This in turn seemed to cause difficulties with place value. Most likely, this problem should either be removed or moved to a later place in the instrument. A much larger group and more questions would be needed to confirm whether, if either of these approaches, would improve reliability.

Table 1*Estimation Attempts and Item Analysis for the Problems in the Interview Activity*

	Sub1	Sub2	Sub3	Sub4	Sub5	Sub6	<i>Item</i> <i>Sum</i>	<i>Difficulty</i> <i>(Mean)</i>	<i>Pearson</i> <i>Correlation</i>
practice	1	2	5	2	2	1	13	2.17	0.63
item 1	2	2	1	1	1	3	10	1.67	-0.80
item 2	1	1	1	3	3	1	10	1.67	0.68
item 3	1	1	1	2	2	1	8	1.33	0.68
<i>Subject</i>									
<i>Sum</i>	5	6	8	8	8	6	41		

The analysis of the interviews resulted in sequences of codes. However, for a given interview segment, different valid interpretations existed. A greatest-common subsequence algorithm (Maier, 1978) implemented in the Unix *diff* utility was used to establish inter-rater agreement. Longest subsequences divided by total codes resulted in 62% agreement across all six interviews. Importantly for this research, a more robust process model was established to describe the process of estimation with the Estimation Calculator (see Figure 7).

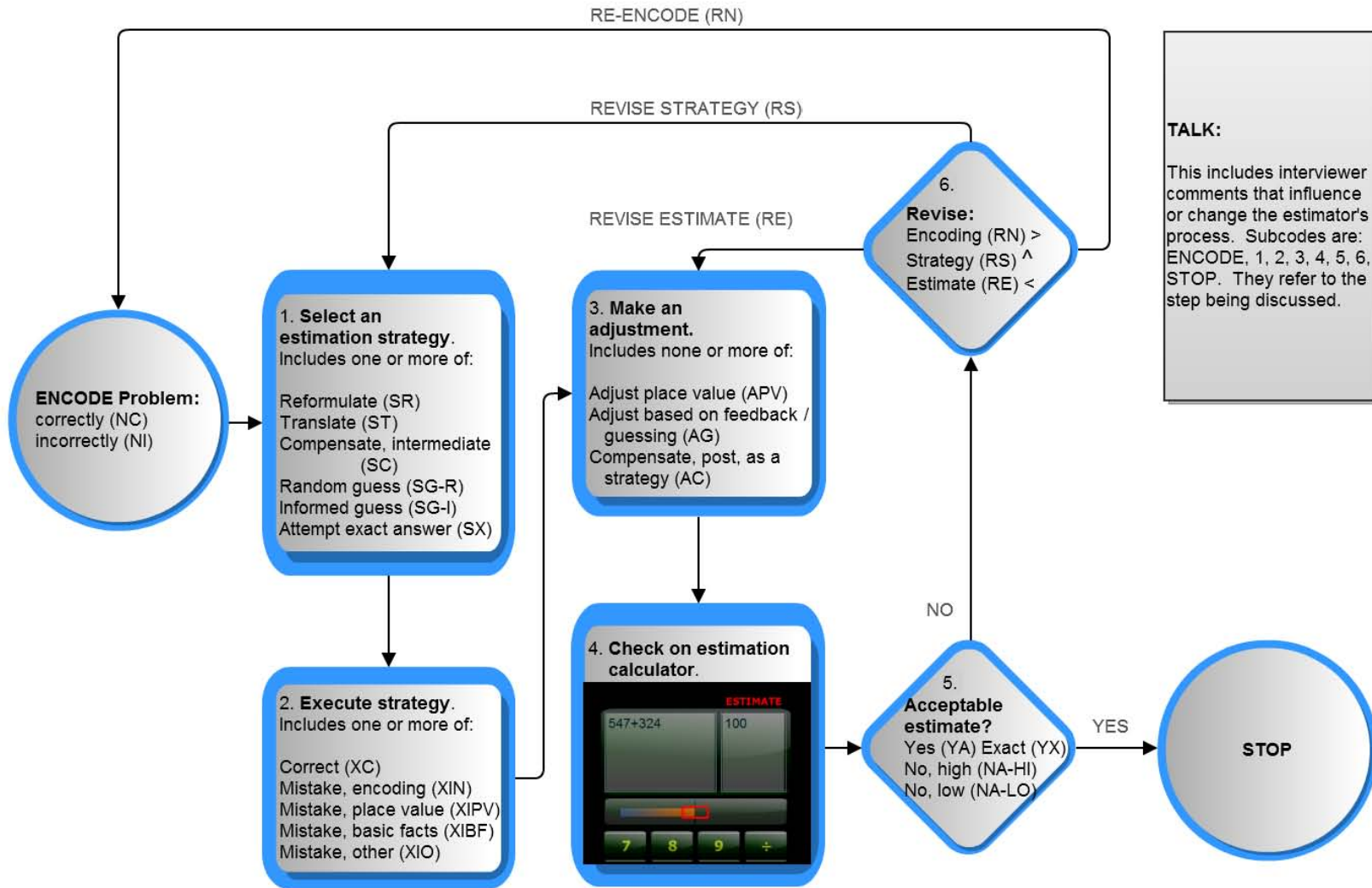


Figure 7: Final Process Model

This new model included all the possible sources of error that could occur as students used the Estimation Calculator. This could include incorrectly encoding the digits or operators, in other words misreading the problem. This occurred on more than one occasion as students responded to problems. Also, the new model distinguishes between two types of guessing, random guessing and informed guessing. Random guessing would include simply choosing an arbitrary number, such as one that led to success on a previous problem (as described in Study 1). Informed guessing is illustrated in the discussion below, for the practice problem, 26×4 . Names are pseudonyms. Paul is the interviewer and Maurice is the student, and Maurice has started the estimation sequence that begins with an estimate of 60:

Paul: "Can you tell me about your original estimate of 60? Tell me the process you went through with those numbers to get 60. Do you remember?"

Maurice: "Well I was just estimating."

Paul: "Ok."

Maurice: "Guess I ... guessing."

Paul: "Well that's perfect. That's what estimation is. But I was wondering if there, like, from the looks – I mean, if the numbers were a million times a million, you wouldn't get 60, right?" (laughs) "So, looking at those numbers, is there a reason that 60 came to mind?"

Maurice: "Well, I thought it was less than a hundred, and 60 is the ... way high, so I just guessed 60."

Maurice seems to be anchoring on the number of digits in the problem. He selects an estimate that has the same number of digits that their product might reasonably have. Perhaps he has performed some form of quick mental calculation, such as 20×4 or 15×4 , however the result and his description suggest an estimate based on magnitude rather than on a specific computation. His choice of 60 is based on the knowledge that it is close to but smaller than 100. Thus this estimate is an example of informed guessing.

Discussion

The revised estimation model (Figure 7) is designed to account for all of the possible sources of error in the process. Thus recorded use of the Estimation Calculator can now be coded in such a way that should provide a description of each stage in the process. The decision point in step 6 of the model now includes a return to the very beginning of the process, to re-encode the problem if a student has read it incorrectly. The model attempts to mimic that of LeFevre et al. (1993) by staying at the level of processes rather than describing specific strategies. However there are steps in the model (2 and 3) where strategy selection would occur. Further research with this model will reveal if it does in fact describe the entire process of estimation with the Estimation Calculator.

Study 3: Volume Estimation

Objectives

The third study addressed a new online tool that extended the use of estimate error feedback shown on the Estimation Calculator display. A new tool, the Spatial Estimator, was developed so that students could estimate the volume of 3D shapes. Students would

estimate aspects of a three-dimensional solid shape shown on the screen. These aspects included length, width, height, and radius for shapes such as cylinders and spheres. The research was designed to address three questions.

The first research question was: How do students use estimate error feedback (EEF) for a volume estimation task? Problem solving involved making measurement estimates using a reference point (Joram et al., 1998) and then synthesizing these estimates through computation estimation. The EEF as error bars were provided for each estimate that a student made (See Figure 8). This included estimates for aspects as well as estimates for total volume.

The shape below is a triangular prism. Remember the volume for a triangular prism:

$$V = L \times (B/2) \times H.$$

The woman next to the prism is 5.5 (5 feet, 6 inches) tall.

The side she is standing next to is the length.

(You may click and drag to rotate the shape.)



Estimate for: height

5

OK



Estimate for: height

3

OK



Estimate for: height

4

OK



Exactly! The actual height is: 4 feet.

Figure 8: Interaction for the Spatial Estimator

Instrument for Dissertation Study – Estimation Questions

Making estimates for aspects of the shapes and making estimates for the volumes suggested two distinct abilities. Measurement estimation and computational estimation have been found to be independent abilities (Hogan & Brezinski, 2003). Estimating aspects, such as length or radius, is a form of measurement estimation (Joram et al., 1998). Estimating volume can incorporate computational estimation (Reys et al., 1982) because it requires the synthesis of the previous aspect estimates combined via geometric formulas. For example, a student might perform measurement estimation on a rectangular prism with dimensions of approximately 3, 4 and 5 feet. Then she might perform computational estimation to estimate 50 cubic feet as the volume of the 3 foot \times 4 foot \times 5 foot prism. She might first reformulate $3 \times 4 \times 5$ to 10×5 because 10 is close to 12.

Following this line of inquiry, the second research question was: Do measurement estimations of one-dimensional aspects and computational estimations of three-dimensional volume represent distinct, separable knowledge components (KCs)? For this, the technique of Learning Factors Analysis was used.

The third research question was: What are the relationships between number of aspect estimates, relative error of aspect estimates, number of volume estimates, and relative error of volume estimates? These relationships were intended to shed light on how strategic or haphazard students' estimation behaviors were. A strong correlation between number of estimates and error of estimates would indicate a haphazard approach.

Participants

Participants included three groups of eighth grade students ($N = 59$) at a central Virginia public middle school. These groups constituted the entire eighth grade, including all levels mathematics of proficiency at the school.

Research Design

The design was exploratory in nature. The intention of the study was to collect descriptive statistics on students' use of the Spatial Estimator, and data for correlation and Learning Factor Analysis. All students were given a forty-five minute regular class period in which to complete the activities. The two eighth grade mathematics teachers devoted their class time to the activities for one day, over the course of three consecutive sections.

Activity

WISE and WISEngineering

The online platform for this research was WISE (<http://wise.berkeley.edu>) although an instance of the WISE server was hosted at a large central Virginia university. The WISE learning management system allows students to work through curriculum step-by-step. Steps are individual web pages based on WISE templates, or step-types. This unit included one step-type that was specifically designed for a new flavor of WISE, WISEngineering (<http://wisengineering.org>). WISEngineering aims to incorporate knowledge integration (Linn et al., 2006) with engineering pedagogy (Chiu & Linn, 2011). Mathematics and science units developed for WISEngineering scaffold students' engineering projects in a systematic and iterative fashion. WISE has been built to support visualizations for teaching science (e.g., Xie & Tinker, 2006). A natural extension for WISEngineering was the incorporation of visualizations for mathematics. The Spatial Estimator uses these new mathematics visualization capabilities to allow students to practice estimation.

The Spatial Estimator

Students use the Spatial Estimator in the manner described above, making estimates for each aspect of a three dimensional shape and then making an overall estimate for the volume of the shape. Students are not expected to know the formulas for the volume of each shape. These formulas are provided in the interface and students use them to make estimates for the volume. A sequence of estimates is shown along the right hand side of Figure 8. The EEF in the form of error bars is shown under each estimate, students are shown the exact answer if their estimate is within the tolerance region. As with the Estimation Calculator, the tolerance region was chosen to be $\pm 15\%$.

Before making estimates with the Spatial Estimator, students took the rotations component of the Purdue Spatial Visualization Test (Guay, 1976). However, the results of this test are beyond the scope of the discussion here.

After completing the mental rotations test, students used the Spatial Estimator to estimate the volume of six shapes. The first shape was a cube. The remaining shapes included two other rectangular prisms, a triangular prism, a cylinder and a sphere. Next to each shape was a representation of a person (See Figure 8). The person's height was given to the students in the explanatory text, so that this could be used as a reference point (Joram et al., 1998). The shapes could be rotated in all three dimensions by clicking and dragging. The position and relative orientation of the person stayed the same as the student rotated the shape. For example, if the student rotated the shape to be upside-down, then the person would also rotate to be shown upside down. The shapes were represented in perspective. The user's ability to rotate the shape to orient it perpendicularly to the view was expected to reduce foreshortening effects (Richardson & Waller, 2005) that can arise with similar estimation tasks in virtual environments.

Instrument for Dissertation Study – Estimation Questions

Each estimate that a student entered was recorded and stored in the database. Students were provided with a space at the bottom of each web-page in the Spatial Estimator to type a description of the processes they used to arrive at their estimate.

Analysis and Results

The first research question was: How do students use estimate error feedback (EEF) for a volume estimation task? An analysis of the distribution of students' estimates was used to address this question. Estimates were grouped by series, where a series of estimates meant all sequential estimates made for the same aspect or volume. Out of 929 series of estimates for both aspects and volume, 100 did not result in a reasonable estimate. In other words, students stopped or gave up before reaching the $\pm 15\%$ tolerance and seeing the correct answer in the output. Patterns of estimation varied widely. On average, students made 2.01 estimates ($SD = 1.67$). However, the distribution of the estimates could be fit very well by an exponential decay curve ($r^2 = 0.96$). The four longest estimate series had lengths of 11, 12, 14 and 17, respectively. More than half (524 series, 56%) contained only one estimate. Of these singletons, 444 estimates (85%) were reasonable.

The second research question was: Do measurement estimations of one-dimensional aspects and computational estimations of three-dimensional volume represent distinct, separable knowledge components (KCs)? Determining distinct KCs for the Spatial Estimator task requires the use of a learning factors analysis (LFA) and the iterative process of determining q-matrices described by Koedinger et al., (2012). The analysis shows whether a smooth learning curve exists for a given KC. Smooth curves mean that the entire set of participants makes fewer errors with each attempt. Data were analyzed using DataShop (<https://pslcdatashop.web.cmu.edu/>). Learning curves were established using only the first estimate in each series. Results are

Instrument for Dissertation Study – Estimation Questions

inconclusive. They fail to show a reduction in errors from problem to problem, revealing improvement within each problem instead. Improvement within a problem is to be expected, since making estimates for width (for example) is easier after having converged on a reasonable estimate for height using the error feedback in the bar.

The third research question was: What are the relationships between number of aspect estimates, relative error of aspect estimates, number of volume estimates, and relative error of volume estimates? Correlations were calculated between these four variables (Table 2). These correlations are intended to shed light on the questions of how systematic students have been in the process of making their estimates. They were calculated per student, and relative error of estimates is calculated as mean absolute deviation (Gorard, 2004).

Table 2

Correlations Among Spatial Estimates and Computational Estimates (n=59)

	NA	EA	NV	EV
Number of Aspect Estimates (NA)		0.082	0.498**	0.072
Error in Aspect Estimates (EA)			-0.101	-0.111
Number of Volume Estimates (NV)				0.446**
Error in Volume Estimates (EV)				

** Correlation is significant at the 0.01 level (2-tailed).

If number of estimates and the mean deviation (error) in those estimates is highly correlated, that would suggest a haphazard estimation strategy, or floundering. It implies that the length of estimate series and the error in those estimate series are linearly related. Students who start with very unreasonable estimates require many attempts to converge on a reasonable estimate. Students who start with moderately unreasonable estimates require a moderate number of attempts, etc. There is a significant correlation ($r = 0.45$, $p < 0.01$) between number of volume estimates and error contained in those volume estimates, suggesting that students were

haphazard in their approach to volume estimation. However, there is no such correlation between the number of and error in aspect estimates ($r = 0.08$). This does not demonstrate conclusively that students were systematic in estimating aspects. However, they may have been able to use cues to converge on reasonable estimates more rapidly than by simply using the error feedback bar.

The other significant correlation that appeared was between the number of aspect estimates and the number of volume estimates ($r = 0.50, p < 0.01$). Thus students who made many attempts for aspects of the shapes also made many attempts at estimating volumes. This runs counter to the hypothesis that these two types of estimation represent the independent abilities (Hogan & Brezinski, 2003) of measurement and computational estimation.

Discussion

Overall, the results highlight the need to interrupt floundering (Anderson et al., 1995) among certain students in order to encourage them to think carefully about their estimates. Most estimate series were fruitful; Of the 829 estimate series that resulted in reasonable estimates, 713 (86%) converged on the reasonable tolerance within three attempts. At the opposite extreme, students made as many as 14 or 17 estimates before providing a reasonable one. The correlational analysis suggests that the students who tended to make these long series of unsuccessful estimates made them in similar proportion for shape aspects and for computed shape volumes. This quantitatively supports the notion of floundering established in Studies 1 and 2. Another result, the indication that students are haphazard in their formation of volume estimates also warrants further attention. Finally, the results of the LFA suggest that much care and attention to instructional design, as well as a larger sample size would be necessary to empirically establish KCs and corresponding problems. The following study (Study 4) included

establishing distinct KCs, but lacked the statistical power to confirm these. A future large-scale study may be able to distinguish and confirm a set of knowledge components necessary for estimation.

Study 4: Estimation in Context

The research described below and in the following two chapters addressed the broad questions, *Is estimation teachable?* and *Is a familiarity with context sufficient for making reasonable estimates for problems in that context?* These two questions cannot be answered by a single research study. Instead, the following more specific questions guided the research:

RQ1: Can instruction on computational and volume estimation combined with automated feedback on estimation in a project-based context lead to students making:

- a) more reasonable computational and volume estimates?
- b) simpler computational and volume estimates?
- c) more nuanced computational estimates?

RQ2: Can an identical project-based context that teaches exact volume computation rather than estimation, and does not include automated feedback lead to students making:

- a) more reasonable computational and volume estimates?
- b) simpler computational and volume estimates?
- c) more nuanced computational estimates?

RQ3) How does reasonableness of estimates differ between students after the two treatments described in R1 and R2 for:

Instrument for Dissertation Study – Estimation Questions

- a) computational estimation problems?
- b) volume estimation tasks that resemble the problems in the unit?
- c) multi-step volume estimation problems that do not resemble the problems in the unit?

RQ4: How do students in the two conditions describe their computational and volume estimation strategies:

- a) before a project-based unit?
- b) after a project-based unit?

Participants

Participants in this study were seventh graders from three sections (hereafter referred to as the three groups; *total* $N = 53$) of mathematics students at a semi-rural, public middle school in central Virginia. Students were assigned into the two treatment conditions so as to maximize the number of students in each treatment. Two of the groups were advanced mathematics sections and the third was a standard-level mathematics class. The larger advanced group ($N = 24$, 63% female) was assigned to the *exact calculation* treatment and so received the exact computation curriculum. They referred to herein as the Large, Advanced, eXact (LAX) group. The smaller advanced group ($N = 14$, 43% female) was assigned to the *estimation* treatment, and are referred to as the Small, Advanced Estimation (SAE) group. The standard-level mathematics group ($N = 15$, 73% female) was taught by the mathematics teacher, who will be referred to simply as "the teacher," with collaborative assistance from a special-education resource teacher. This group also received the *estimation* treatment and is hereafter called the COllaborative Estimation group (COE).

Instrument for Dissertation Study – Estimation Questions

At the time of this writing, the most recent demographic information for the middle school was available on the school's website (WMS Fact Sheet, 2013). According to this data, approximately 14 percent of the student body is African-American, 6 percent is Hispanic, 77 percent is white, and 3 percent represents other ethnicities or are unreported. Roughly 40 percent of students receive federally funded free or reduced-price lunch, 2 percent have limited English proficiency, 13 percent have disabilities, and 13 percent have been identified as gifted.

The concern of equity was addressed in this research by providing both treatments with access to the desktop manufacturing technology. This was anticipated to have a motivating and engaging effect on the students (Berry et al., 2010). The estimation groups were exposed to activities with more feedback based on estimation-errors, while the exact-calculation group was given activities based on the requirement of exact answers. Attention was paid to making both treatments as similar as possible while maintaining the key empirical distinctions.

Research Design

This quasi-experimental study introduced the new independent variables of estimation instruction and context. The independent variable of feedback from the previous studies was under investigation here as well. The two treatments were the estimation group and the exact calculation group. The estimation group received instruction on strategies for estimation based on the principles outlined here. That group made use of the Estimation Calculator and a new estimation practice step-type in WISEngineering throughout a five-day unit. The exact group had the same amount of time to study 3D geometry in the same context. However their work focused on making exact volume calculations rather than estimates.

Instrument for Dissertation Study – Estimation Questions

Activity

The overall structure of the activities and a data collection strategy is shown in Table 3 and a breakdown of the curriculum for both treatments is shown in Table 4.

Table 3

<i>Data Collection Strategy</i>		
	Estimation Group	Exact Group
Pre-Assessment	Computational Estimation Problems	Reformulated / Translated Problems
	Volume Estimation Problems	Reference Points or Other Anchors
	Transfer Volume Estimation Problems	Estimates
	Interviews	Student Descriptions of Strategies
Intervention	Context: Design and Digital Fabrication	
	CAD Software Lessons	
	Estimation Lessons	Exact Computation Lessons
	Software-Based Feedback for Estimation	
Post-Assessment	(same as pre-assessment)	(same as pre-assessment)

Assessment Instruments

Pre- and post-assessments were identical. An online assessment instrument was developed that collected three kinds of information on student estimates: (1) students' reformulated or translated problems; (2) students' reference points or other anchors; and (3) students' estimates. The instrument included both Computational Estimation (CE) problems and Volume Estimation (VE) problems.

Figure 9 shows part of the instruction sheet that participants received on how to answer the assessment items for estimation. This includes a screenshot of the interface with an overlay

Instrument for Dissertation Study – Estimation Questions

describing the four text areas on the screen where users can type the descriptions of their strategies as well as their final estimates. The top-left area (marked **1. WORDS**) prompts users to provide descriptions in words, to elicit the anchor, reference point or strategy used. The top-right text area (marked **2. NUMBERS / SYMBOLS**) asks users to write their strategies with numbers and symbols, to elicit the reformulated or translated version of the problem. This region also includes a number-pad with all of the numbers and symbols that users can use. This text-area is validated so that symbols absent from the number-pad (such as letters) are not accepted and incomplete mathematical expressions are flagged with an "Incomplete answer" warning that disappears when students complete the expression. Toward the bottom of the figure is the text-area (marked **3. ADJUSTMENT (optional)**) for a user's adjustment such as a place-value adjustment or the use of a compensation strategy. This area remained invisible unless a user clicked on a checkbox with the label, "Check this box if you want to adjust your estimate." Finally, the bottom of the figure (marked **4. ESTIMATE**) elicits the user's final estimate. Problems for volume estimation were similar, although the prompts for users to write their estimates differed from those for computational estimation (See Figure 10).

Instrument for Dissertation Study – Estimation Questions

QUIZ QUESTION

Make an estimate for the problem:

$$23 \times 78$$

Describe your estimation strategy or strategies with **words** here:

round 23 down. round 78 up.

Describe your estimation strategy or strategies with **numbers and symbols** here:

$\approx 20 \times 80$
 $\approx 2 \times 8$
 $= 16$

OK

7	8	9	÷
4	5	6	×
1	2	3	-
0	.	()	+
≈	=	←Bsp	Clr

1. WORDS

2. NUMBERS / SYMBOLS

3. ADJUSTMENT (optional)

4. ESTIMATE

☒ Check this box if you want to adjust your estimate.

Describe how you adjust your estimate here:

put two 0s at the end.

Use this area to write your final estimate:

1600

Figure 9: The estimation assessment interface for computational estimation

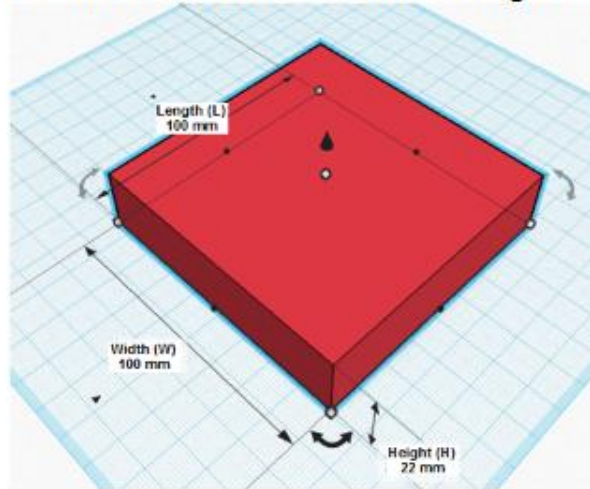
Instrument for Dissertation Study – Estimation Questions

QUIZ QUESTION

Make an estimate for the problem:

Estimate the volume of this rectangular prism. The volume for a rectangular prism is $V = L \times W \times H$.

Length = 100mm. Width = 100mm. Height = 22mm.



Describe your estimation strategy or strategies with **words** here:

Describe your estimation strategy or strategies with **numbers and symbols** here:

Type your answer or click on the number-pad here:



☐ Check this box if you want to adjust your estimate.

Use this area to write your final estimate:

Figure 10: The assessment interface for volume estimation

Computational estimation problems included one based on each of the four standard functions (+, −, ×, ÷). Problems were taken from the ACE Test (Reys et al., 1980) and from problems that provided useful results in the previous studies. The volume estimation problems

Instrument for Dissertation Study – Estimation Questions

required analysis of figures that resembled the students' CAD designs. Because of the nature of desktop manufacturing and the software that was to be used, these figures were no larger than 10cm cubed and used to be combinations of three-dimensional solids usually covered in upper-elementary grades (National Governor's Association, 2010a).

Interviews also provided data on student strategies. Students were asked to perform talk-throughs of their estimation strategies for computational and volume estimation problems. As in Study 2, the interviewer made an effort to refrain from suggesting anything to the students. Two students were chosen at random from each group to be interviewed as part of the pre- and post-assessment. A transfer volume estimation problem was added to the post-interview. This problem required students to make a multi-step volume estimate in a context similar to the one in which students practiced with in the classroom intervention. This item involved estimating of the volume of an object that had been 3D printed, and required students to choose reasonable units for estimates of volume and printing time.

Assessment Development

Assessment instruments were developed by the lead researcher, with assistance from an expert in mathematics education. The instrument was heavily influenced by the ACE Test (Reys et al, 1980), which includes 4 segments:

I: Straight Computation

II: Application

III: Calculator Segment

IV: Attitude / Concept Questions

Questions in all four sections of the ACE are challenging, because it was originally designed for use with participants ranging in age from children to adults. Therefore,

Instrument for Dissertation Study – Estimation Questions

modifications were made to items in the development of the assessment instrument (See Appendix A for a full list of the problems and modifications that were made to the problems.) Straight Computation Estimation problems in this set borrowed heavily from questions from sections I and IV of the ACE test. Problems that had provided useful insights into student strategies during the previous three studies were also used. One example of a modification to the original problems in the ACE test was a change from the original problem, $42,163 \div 73$, to the simpler problem, $426 \div 7$. The rationale for this kind of change was that the challenges to estimation, particularly division estimation, center on manipulation of the front-end digits. To address more fundamental knowledge of mathematics facts and number sense, while reducing cognitive load of digit manipulation, problems such as this one were reduced by several orders of magnitude while retaining the one or two most significant digits.

For the assessment segment on Volume Estimation, original problems were created using TinkerCAD, the 3D software that was part of the curriculum, although screenshots from the software program were modified to exclude decimal portions, which in each case were zeroes (.00), and to include millimeter (mm) unit labels and dimension names (e.g., Height). These problems were designed to increase in difficulty. The first estimation problem involves estimating the volume of a cube that is 10 mm on each side, to address the fundamental ability to work with the $V=L \times W \times H$ formula. These assessment items increased in difficulty, so that the last of seven volume estimation problems was a multi-step problem, requiring the subtraction of a rectangular prism shape (hole) from an outer cylinder.

Curriculum Development

The lead researcher created the five-day curriculum for each group, using materials from previous estimation lessons and research studies. Development of the curriculum centered on

Instrument for Dissertation Study – Estimation Questions

translating the outline shown in Table 4 into estimation and exact calculations lesson plans². The lead researcher met twice with the teacher before the intervention to lay groundwork for the research study.

Table 4

Activities for the two treatment groups

Estimation Group	Exact Group
<u>Phase 1: Computational Estimation</u>	<u>Phase 1: Solid Geometry</u>
Part 1.1: Discussion of Importance of & Reasons for Estimation	Part 1.1: Discussion of Formulas for Solids
Part 1.2: Practice with Computational Estimation Strategies	Part 1.2: Practice Calculating Volumes of Solids
<u>Phase 2: Solid Geometry</u>	<u>Phase 2: Multi-Step Geometry (Composite Shapes)</u>
Part 2.1: Discussion of Formulas for Solids	Part 2.1: Discussion of Volume for Composite Shapes
Part 2.2: Practice Calculating Volume of Solids	Part 2.2: Practice Calculating Volume of Solids
<u>Phase 3: CAD Design</u>	<u>Phase 3: CAD Design</u>
Part 3.1: Experimenter Demonstration of 3D Printing	Part 3.1: Experimenter Demonstration of 3D Printing
Part 3.2: Using <u>TinkerCAD</u>	Part 3.2 : Using <u>TinkerCAD</u>
<u>Phase 4: Estimating Composite Shape Volumes</u>	<u>Phase 4: More Calculating Composite Shapes</u>
Part 4.1: Discussion of Volume of Composite Shapes	Part 4.1: Recap Discussion of Volume of Composite Shapes
Part 4.2: Practice Computationally Estimating the Volume of Composite Shapes	Part 4.2: Practice Calculating the Volume of Composite Shapes
Part 4.3: Practice Estimating the Volume of Composite Shapes using Reference Points	Part 4.3: Each Student Calculates the Volume of the Other Student Designs
Part 4.4: Each Student Estimates the Volume of the Other Student Designs	
<u>Phase 5: Production</u>	<u>Phase 5: Production</u>
Part 5.1: Printing the 3D Design of One Student at Random	Part 5.1: Printing the 3D Design of One Student at Random
Part 5.2: Printing the 3D Design of the Best Estimate for the First Design	Part 5.2: Printing the 3D Design of Another Student at Random

² Complete lesson plans are available online at

Instrument for Dissertation Study – Estimation Questions

The special-education resource teacher who collaboratively taught the COE group was present for the second meeting, and provided suggestions as well. The simplifications to the ACE instrument, described above, were also the result of the second meeting, which included both teachers.

Software Development

Software was developed and tested for WISEngineering by the lead researcher. One of the other programmers on WISEngineering helped with the integration of new software into the existing WISEngineering servers. This involved two new step-types for WISEngineering, one of which was used for the assessment portion of the study, and one of which was used for the interactive practice sessions. Both step-types resembled Figures 9, and 10, although the interactive step-type hid the response shown, instead revealing a box for estimates and a feedback bar. After the third unreasonable estimate (outside the $\pm 15\%$ region), the response screens appeared, prompting students for the four inputs shown in Figures 9 and 10.

Intervention

The intervention consisted of five phases for each of the two treatment conditions (all three groups). These are shown in detail in Table 4. The majority of the classroom activity was performed on computers. The entire intervention is described below as a five-phase sequence. This roughly corresponded to a five-day unit, although the COE group split their use of the 3D design software over two days, and the timing of the assessment corresponded with the final day for all three groups.

Instrument for Dissertation Study – Estimation Questions

The phases of Solid Geometry, CAD Design, and Production were part of the treatment for both groups. Solid geometry lessons involved reviewing and practicing with the formulas of a few basic shapes, including rectangular prisms and cylinders. The CAD Design phase required the students to learn the basics of a CAD software tool (TinkerCAD, <https://tinkercad.com>) that allows for adding and subtracting shapes. The design centered on the simple design activity of designing a pencil-holder. After completing the pencil holder, students were free to design their own object (Gershenfeld, 2005) to allow them imaginative freedom. Production was the last phase. This involved the actual production, via a 3D additive printer, of two student designs from each group. All six of these were chosen by lottery. A worksheet was designed as part of Phase 4 for students in the estimation groups to estimate, and for the LAX groups to calculate, the volume of two pencil holders that had already been printed in class. This was intended to highlight the relevance of making estimates based on what is not known exactly beforehand (Usiskin, 1986).

The estimation group received instruction on estimation. This included a class discussion on its importance, situations where it is needed, and lessons on strategies for improving estimates. In particular, strategies were taught for multiplication estimation and compensation for addition, subtraction and multiplication (See Appendices MULT_EST_WORKSHEET and COMPENSATION_WORKSHEET). Student had the chance to practice with the Estimation Calculator to reinforce the use of these strategies. The estimation treatment groups spent one day on estimation of volume problems.

Meanwhile, the exact treatment group spent this time practicing finding exact solutions to multi-step geometry volume problems. This group spent one day working this type of problem

Instrument for Dissertation Study – Estimation Questions

and one day doing an extension activity on the features of geometric solid figures on an interactive website.

Before designing on CAD software, all groups spent part of one lesson learning the formula to estimate necessary print-time in minutes based on the volume, in cubic millimeters, of a 3D design. Thus students' volume estimates, multiplied by some ratio of volume to time needed to digitally fabricate a design, provided an estimate of how long the process might take.

Analysis

All four of the research questions were addressed via an analysis of the four sources of data:

1. Students' Reformulated or Translated Problems
2. Students' Reference Points or Other Anchors (self-reported)
3. Students' Estimates
4. Students' Descriptions of Their Estimation Strategies

These data sources are shown in the right-hand column of Table 3

Research question 1 is addressed by including the four independent variables in the estimation treatment group:

1. Instruction
2. Reasonable Tolerance for a Problem
3. Problem Context
4. Feedback

All three parts of RQ1 are addressed based on an analysis of computational and volume estimates in pre- and post-assessment instrument. Dependent measures were derived from

Instrument for Dissertation Study – Estimation Questions

estimates and additional information provided (Figure 9). Simplicity was calculated by way of complexity, measured in terms of significant digits retained in reformulation. For example, the number 352 can be rounded to 400, which would be scored as having a complexity of 1, because it retains 1 significant digit. Likewise, if 352 is truncated to 300, it has a complexity of 1.

Rounding 352 to 350 would result in a complexity of 2 for containing 2 significant digits, etc.

Lower complexity scores generally mean simpler estimates. This analysis ignores the magnitude of the original operands and some cognitive load components such as rounding up or rounding down (Lemaire & Lecacheur, 2011), but provides a rough approximation of estimation difficulty.

Significant digits were summed for a total complexity score. Complete translation of a problem was treated in the same way. For example, translating $(48 + 49 + 52)$ into (50×3) results in a score of 2 for the two remaining significant digits: the 5 and the 3.

Reasonableness of estimates can be calculated directly as mean absolute deviation (Gorard, 2004) of estimate errors. For this study, in order to standardize estimates to the same scale, the following transformation was used:

$$T_i = \log_{10} \left(\frac{E_i}{X_i} \right)$$

where E_i was an estimate for problem i , X_i is the exact answer, and T_i is the transformed estimate. Characteristics of this transformation are discussed in the following chapter.

Nuance is too ambiguous to analyze in the context multi-step volume estimates. Deriving nuance for computational estimates involved summing altered digits, as did the measure of complexity just described. It also takes altered operators into account. The following algorithm will be used: Each digit in the reformulated problem that differs from a digit that would be obtained by algorithmic rounding will be added to the nuance score. For example, algorithmic

Instrument for Dissertation Study – Estimation Questions

rounding would generally involve reformulating 62×23 as 60×20 . If instead, a student rounds to 60×25 , this reformulation received a nuance score of 1 because final digit (5) in 25 differed from the final digit (0) in 20. This was intended to capture intermediate compensation (Reys et al., 1982).

Problems that students thoroughly translate, such as converting 62×23 into 600×2 or 6×2 would have received one nuance point for every single in digit in their new problem plus one for any new operator. This adjustment to nuance score was intended to capture translation (Reys et al., 1982). However, student data did not reveal any use of this kind of translation strategy.

Research Question 2 was addressed by including only the Problem Context independent variable in the exact treatment group. Instruction on estimation and estimate feedback was excluded. The remaining time in the intervention was spent practicing exact computation strategies for multi-step volume problems. Dependent variables were derived from the pre- and post-assessment instrument and analyzed in the same manner as for the estimation treatment group.

Research Question 3 required a 3×2 (group by pre-post) ANOVA, with mean deviation of estimates averaged over the three tasks as the dependent variable. Subsequent pair-wise comparisons were performed on each of the three tasks, with a Bonferroni correction (Lomax, 2007) on the alpha cutoff to compensate for increased likelihood of Type II errors.

Research Question 4 was designed to provide a qualitative support to the quantitative analyses described above. Interviews were conducted with two students from each group. Each interview consisted of computational estimation problems and multi-step volume estimation problems solved as a think-aloud by the student. Interviews were analyzed to reveal insights that students spontaneously discussed while making estimates.

Instrument for Dissertation Study – Estimation Questions

Chapter 4: Results

Introduction

This chapter describes the results, starting with the quantitative results from the assessment activities, continuing with results from the interview segment with students as well as an additional follow-up interview that was held with the teacher, and concluding with a discussion of some of the more prevalent sources of estimation error. The results from the assessment activities are structured to address RQ3 first, since this question concerned a test of pre- to post- differences in the reasonableness of estimates for all three groups. This omnibus test (repeated-measures ANOVA) is a logical entry point for analysis. Next, results concerning the simplicity and nuance of student responses, from RQ1 and RQ2 are presented. The theme of volume estimation with a cylinder was salient, and is featured in the subsequent section. After that follow the results from the student interviews (RQ4), a discussion of students' resulting 3D designs, as well as quotes from the follow-up interview with the teacher. Finally, examples of outlying estimates and flawed student strategies are shown. All of these results are summarized in a table at the beginning of Chapter 5.

Analysis Techniques

Results were collected in MySQL database connected to the WISEngineering server. These were downloaded from the server machine and responses to the relevant step-types were filtered through a secondary server owned by the lead researcher. Analyses were performed using Excel 2010, and SPSS, Version 20.

Instrument for Dissertation Study – Estimation Questions

Student responses to estimation provide a wealth of information. For this analysis, the "first pass" consisted of looking at distributions of all student estimates from all groups both before and after the intervention., which tend to cluster around the exact answer, with some estimates falling above and some falling below. In order to place underestimates and overestimates on the same scale, estimates were transformed using the following function:

$$T_i = \log_{10} \left(\frac{E_i}{X_i} \right)$$

Here T_i represents the transformed estimate to problem i , X_i represents the exact answer to problem i , and E_i represents a students' estimate for that problem. The benefit of this transformation is that it makes order-of-magnitude errors very apparent. For example, if the exact answer to a problem is 100, then an estimate of 10 will be transformed to -1, and an estimate of 1,000 will be transformed to 1. Put in terms of student estimation strategies, this transformation gives a good indication of how many zeroes a student added unnecessarily or left off. In so doing, it gives order-of-magnitude overestimates and underestimates roughly the same weight³.

Altogether, there were 867 estimates given by students for the 18 computational and volume estimation problems. These included four problems at the beginning of the assessment instrument, testing each of the four basic mathematical functions (+, −, ×, ÷) as well as seven geometry problems (see Appendix A), and additional questions involving the use of percentages

³ Note that close underestimates and close overestimates result in transformed scores with different absolute values. For example, an estimate that is 15% below the correct answer becomes -0.07, while an estimate that is 15% above becomes 0.06. Thus the logarithmic transformation for "punishes" underestimates slightly.

and fractions. A histogram of all students' estimates, including pre- and post-intervention assessment reveals modes near -1 and 1, as well as a large cluster close to the exact answer (see Figure HISTO). This excludes four estimates that were deemed outliers because their transformed error was greater than 4.0, so the figure includes only 863 responses.

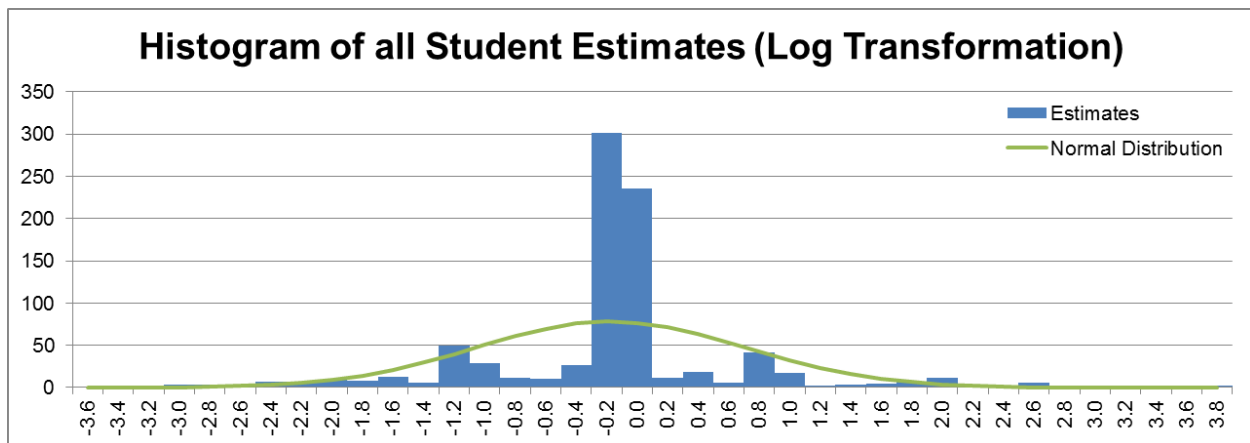


Figure HISTO: A histogram of all student estimates, shown with a base-10 logarithmic transformation. $N = 863$ responses.

This multimodal distribution indicates that over half of estimates fall close to the exact answer. However, estimates also cluster near the negative and positive integers, particularly near -1.0 and 1.0. This fits with the premise that students will often provide answers that differ from the exact answer in magnitude. The distribution is also not symmetric around zero. The bin-widths in this histogram are 0.2, which corresponds to a range of -37% and 58% estimate-error for the left and right bins straddling the 0.0 mark, respectively. More estimates fall just below the exact answer than just above it. For example, an estimate of 1,200 for 62×23 can be explained by a reformulation of 60×20 and has an error of -16%. Its transformed error is just below -0.07. Meanwhile, a similar estimate that lacks a zero, 120, has a transformed error just below -1.07. Thus, the estimates that cluster just *below* the integers (-1.0, 0.0, 1.0, etc.) are estimates that likely used a rounding-down or truncation strategy.

Departure from Normality

The mean transformed error in the above distribution ($M = -0.07$, $SD = 0.88$) corresponds to an average error of just below -15%, falling outside the acceptable tolerance on the Estimation Calculator. This includes estimates from before and after the intervention. Unfortunately, this distribution varies greatly from the normal distribution, $\chi^2(35, N = 863) = 1587.16, p < .001$. Similarly non-normal distributions were found for subsets of the data, split by problem type, group and time.

Limited Problem Set

The last seven problems in the assessment instrument were considered to be additional problems. These questions were designed to be reachable by only the fastest students, and received only 21 responses from all 53 students during the pre-assessment. They are the MORE-COMP questions at the end of Appendix A. They addressed use of decimals and fractions, which students were not directly taught during the interventions. The following results are drawn from a subset of the data that excluded these extra problems, thus including only the four basic function questions and the seven geometry questions. The distribution of the remaining 11 questions is non-normal, $\chi^2(33, N = 767) = 1848.06, p < .001$. Descriptive statistics indicated a lower mean than described above and a similar standard deviation ($M = -0.13$, $SD = 0.83$). Visual inspection of the distribution suggested that the low mean was due to a concentration of underestimates (See Figure HISTO_11_PROBS). This mean of -0.13 corresponds to an estimate with -26% error.

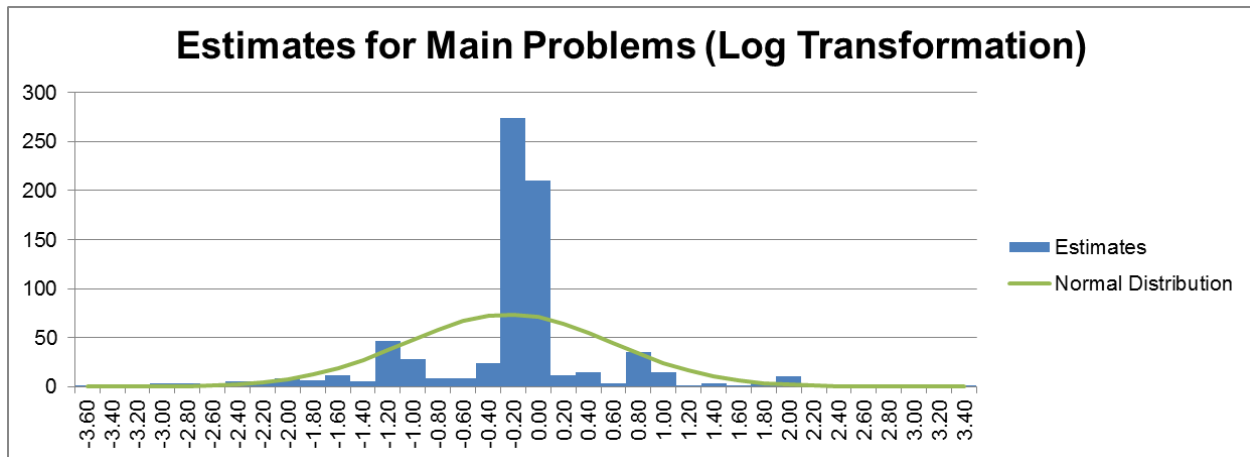


Figure HISTO_11_PROBS: A histogram of a subset student estimates from the main 11 problems in the assessment, shown with a base-10 logarithmic transformation. N of 767 responses.

Because of the non-normality of the data, further transformations were applied to answer the research questions regarding changes to the reasonableness. The algorithms defined in the previous chapter were used to establish simplicity and nuance of student estimates from pre- to post-intervention.

Reasonableness of Estimates Across All Groups

Research Question 3 concerns repeated measures changes across groups. Since it requires an omnibus analysis, it is the most logical question to address first. The non-normality mentioned above is investigated further in the next section by measures of skewness and kurtosis.

Reasonableness Analysis

Data for the computational estimation (CE) and volume estimation (VE) problems were analyzed separately. First, absolute values were taken to determine the magnitude distance of an estimate from no error. Four student responses appeared as apparent outliers and were

Instrument for Dissertation Study – Estimation Questions

eliminated from the analysis because they distorted the distribution of the estimates. Then, this absolute magnitude error was averaged across problems by student, resulting in a mean absolute magnitude error (MAME) score per student.

Computational Estimation

Figure FIGURE 11, below, shows the distributions of students' estimates. Clearly the post-assessment elicited more estimates that were 3 and 4 orders of magnitude from the exact answer from all three groups. The analysis below clarifies the movement of the median values, which are all within or very close to the reasonable range of $<.07$.

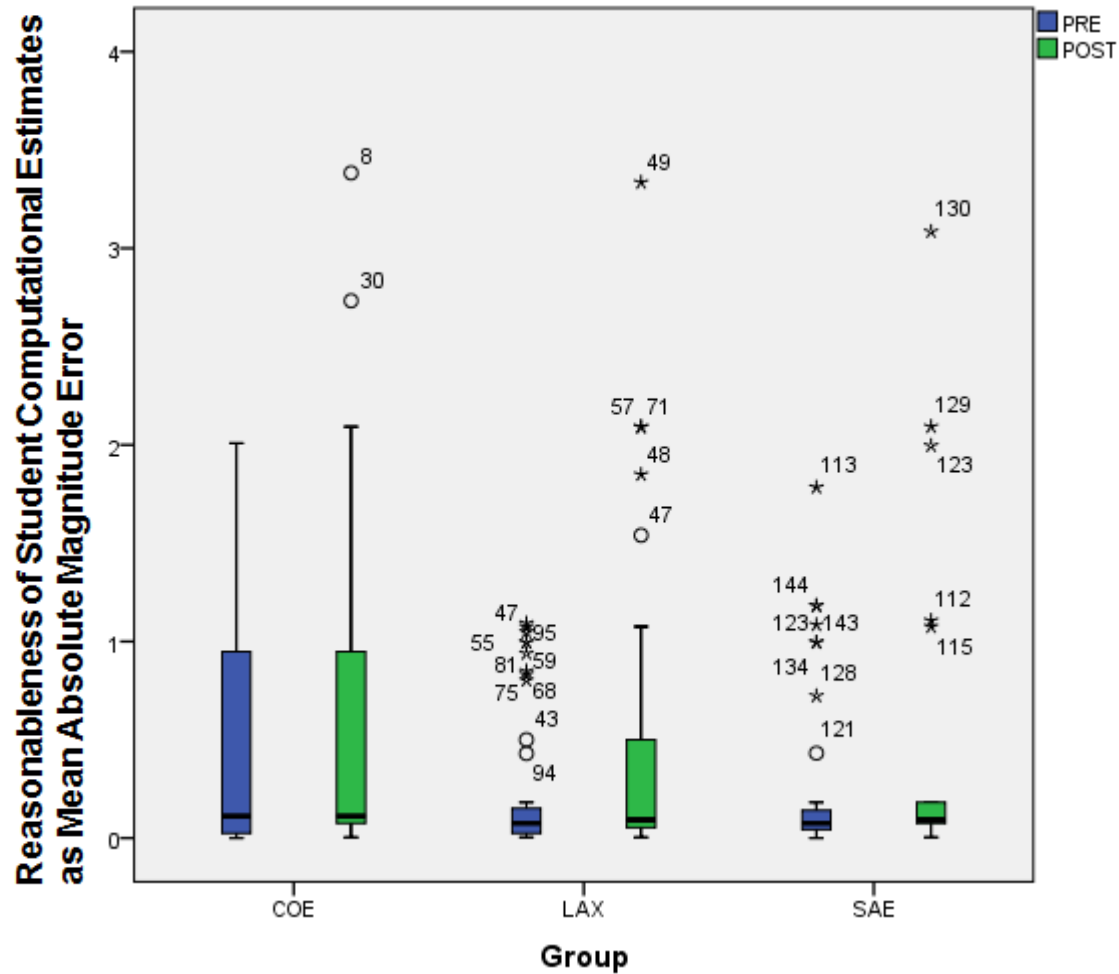


Figure 11: Boxplot of the MAME of students' Computational Estimates

- Moderate outliers
- * Extreme outliers

Table 5

Medians and Inter-quartile Ranges for the Reasonableness of Computational Estimates

	PRE		POST	
	Median	IQR	Median	IQR
COE	0.11	0.94	0.11	0.89
LAX	0.07	0.13	0.09	0.47
SAE	0.07	0.11	0.09	0.11

Analysis of skewness and kurtosis of these scores is summarized in the following table:

Table 6

Descriptive Statistics for the Reasonableness of Computational Estimation (CE) and Volume Estimation (VE) Problems

		Mean	Standard Deviation	Skewness	Kurtosis	Shapiro- Wilk's test (<i>p</i>)
PRE	CE Problems	0.30	0.21	0.53	-0.47	.02
	VE Problems	0.57	0.59	1.49	1.54	<.01
POST	CE Problems	0.44	0.47	1.79	3.02	<.01
	VE Problems	0.45	0.52	2.15	5.14	<.01

According to Kline (1998), a skew values less than 3 and Kurtosis values less than 10 indicate that data do not exhibit extreme non-normality, and by that measure these data do not. However, Shapiro-Wilk's tests were also performed, resulting in the significance values shown at the right side of the table above. Each one of these tests shows a significant deviation from normality. Therefore, as a next step, data were ranked across students in all three groups, and a Van der Waerden transformation was applied so that the data would be distributed normally. After the adjustments described above were applied to the data, a repeated measures analysis of variance (ANOVA) was performed. Missing data accounts for the degrees of freedom being below 50.

Table 7

ANOVA for the Reasonableness of Computational Estimates, after Van der Waerden transformation

	<i>df</i>	<i>MS</i>	<i>F</i>
Group	2	4.77	4.43*
Time	1	<.01	< .01
Group \times Time	2	0.22	.44
Between Group Error	39	1.08	
Within / Time Error	39	0.49	

* indicates significant at the $p < .05$ level

An examination of the transformed means by group revealed that the COE group made more unreasonable estimates pre- (0.59) and post-intervention (0.38) than did the SAE group ($pre = -0.35$, $post = -0.32$), or the LAX group ($pre = 0.76$, $post = 0.88$). A measure of the effect size of this relationship was also obtained ($partial \omega^2 = .02$). The difference between groups is visible in the following figure. As this analysis reveals, the COE group differed from the other two groups. Students in the COE group had less of a concept of reasonable computational estimation before the unit than the other two groups, and a slightly, though not significantly, greater grasp of it afterward.

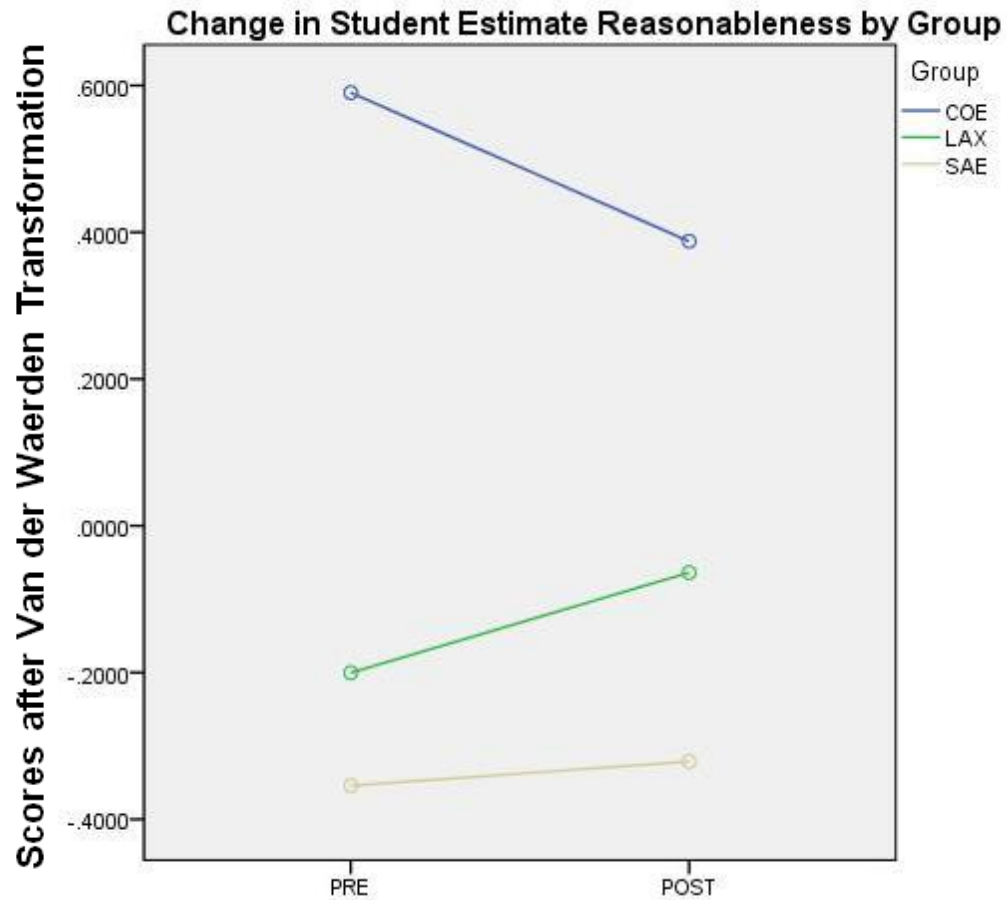


Figure 12: Changes in mean student Computational Estimate reasonableness from pre-assessment to post-assessment. Data shown are transformed MAME scores.

Note that this is a measure of errors in estimates, so the desired result is a decrease. More reasonable estimates have lower error.

Volume Estimation

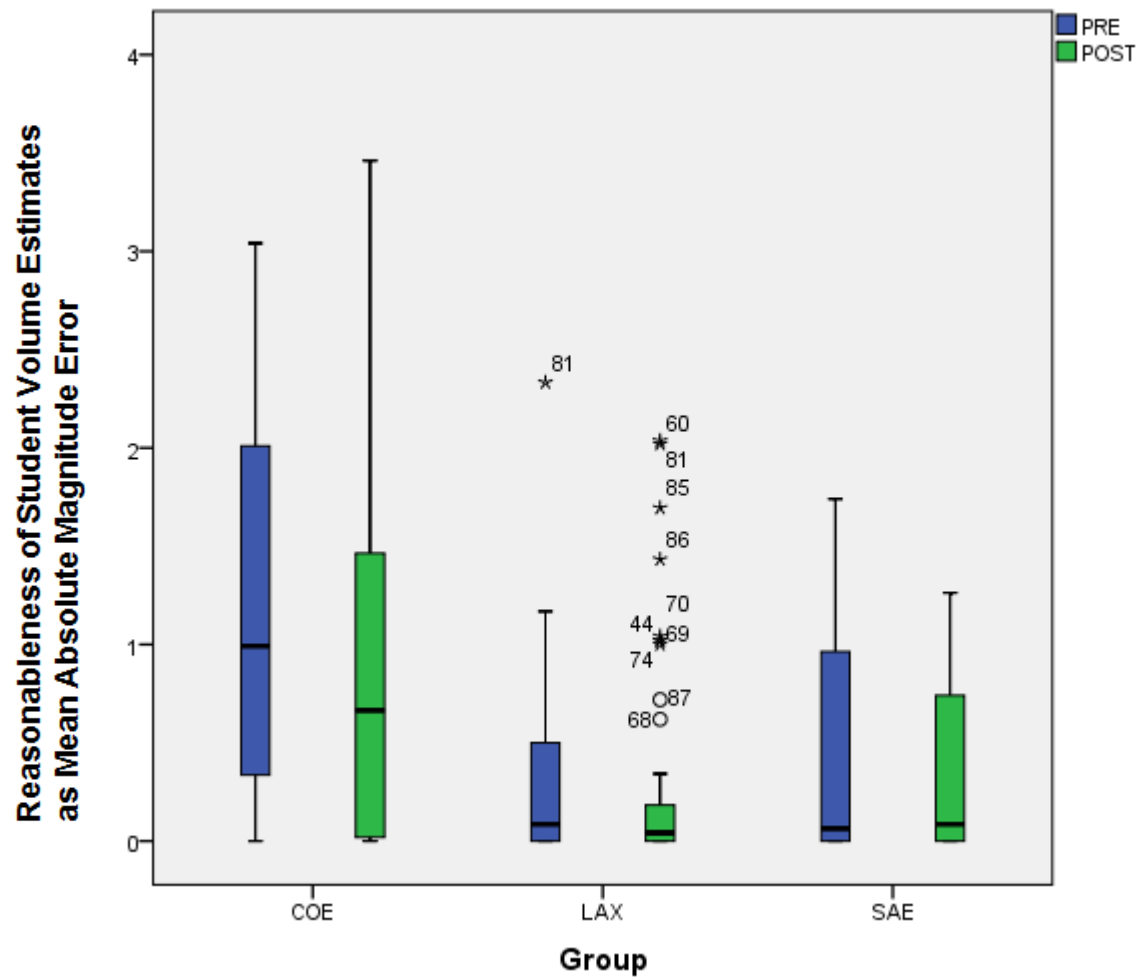


Figure 13: Boxplot of the MAME of students' Volume Estimates

- o Moderate outliers
- * Extreme outliers

Table 8

Instrument for Dissertation Study – Estimation Questions

Medians and Inter-quartile Ranges for the Reasonableness of Volume Estimates

	PRE		POST	
	Median	IQR	Median	IQR
COE	0.99	1.68	0.66	1.47
LAX	0.09	0.50	0.04	0.18
SAE	0.06	0.97	0.09	0.79

Table 9

ANOVA for the Reasonableness of Volume Estimates, after Van der Waerden transformation

	<i>df</i>	<i>MS</i>	<i>F</i>
Group	2	3.71	3.05
Time	1	0.09	0.28
Group × Time	2	0.70	2.32
Between Group Error	34	1.22	
Within / Time Error	34	0.30	

Parentheses indicate Mean Squared Error.

* indicates significant at the .05 level. (There were no significant effects.)

There were no significant results that occurred between students' estimates for volume estimation problems between groups or within-students from pre-to post assessment. As can be seen in the following figure (Figure 14) the COE students have made estimates that are more reasonable, by comparison with the other groups and their original estimates.

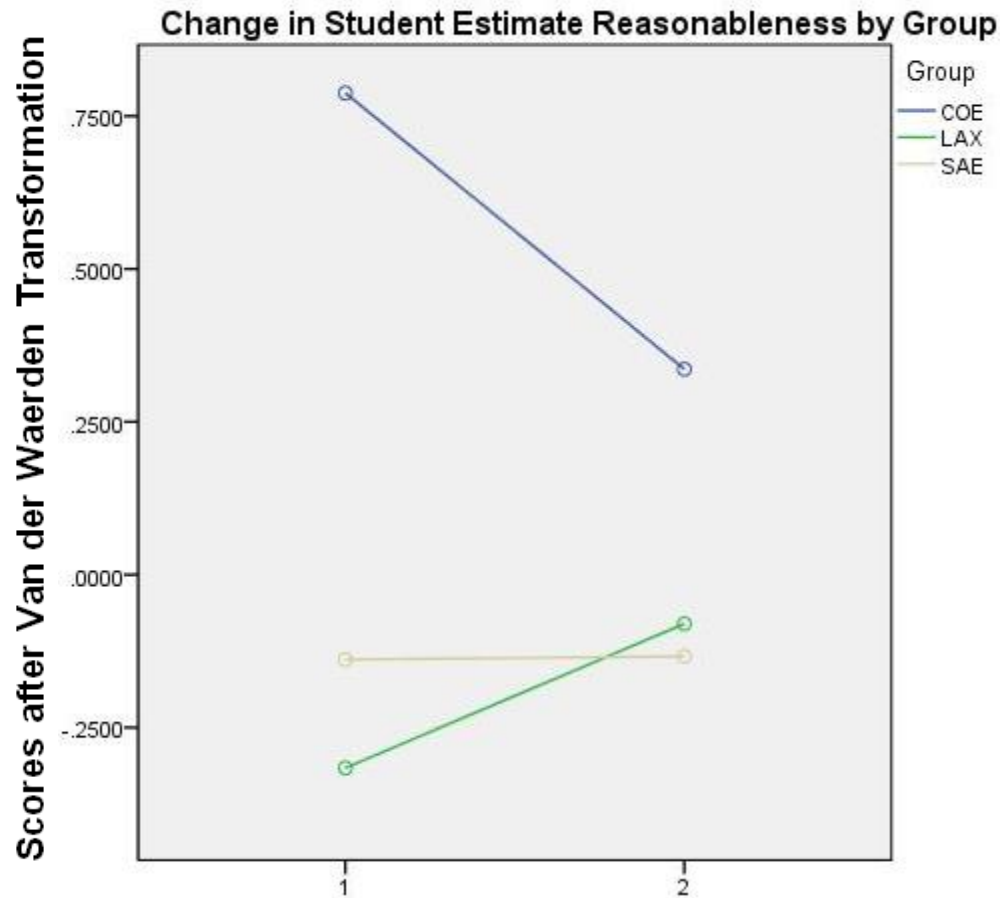


Figure 14: Changes in mean student Volume Estimate reasonableness from pre-assessment to post-assessment. Data shown are transformed MAME scores.

Only one multi-step estimation problem was included in the assessment tool. This is the Composite item, which is number 7 in Appendix A. It had a 17% and a 23% response rate on the pre- and post-assessments, respectively. Its responses are included in the previous analysis, and are not analyzed separately because of the low response rate. However, one multi-step problem was included in the post-interview. Student responses to this item are discussed in the interview section below.

Simplicity and Nuance for the Estimation Groups

Research Question 1 asks whether instruction on computational and volume estimation, combined with automated feedback on estimation in a project-based context, can lead to students making more nuanced computational or volume estimates.

Reasonableness

Since there was no main effect on students' estimates for time, no follow-up comparisons are necessary here. As can be seen in figure FIGURE 11, change in error in students' estimates is modest, and did not reach significance.

Simplicity

Simplicity Algorithm

The simplicity algorithm is based on counting the digits used by a student in his or her reformulation of the problem. Many students created a multi-step reformulation, and thus the complexity score was calculated using the last reformulation in the sequence. For example, one student wrote the following sequence, for the problem 62×23 :

$\approx 60 \times 20$

$\approx 6 \times 2$

This includes two reformulations. The second one is the final formulation, and is therefore used as the reformulation for the complexity measure. In fact, these are both considered to have equivalent complexity because trailing zeroes are ignored. For multi-step reformulations, the reformulation was considered to be one that combined all of the operands that were used in the problem. For example, a problem requiring that the student estimated the volume of an elliptic cylinder, elicited the following sequence of reformulations:

Instrument for Dissertation Study – Estimation Questions

$$3 \times 10$$

$$\approx 30 \times 20$$

$$\approx 600 \times 10$$

$$\approx 6000$$

This is a piecewise reformulation of $3 \times 10 \times 20 \times 10$, so this is considered to be the student's actual reformulation.

Pre- to Post-Assessment Comparison

For the computational estimation problems, students' response rates varied greatly. All items for which students did not provide information on their reformulation strategy were eliminated, leaving the following numbers of student responses:

Table 10

*Overall Response Rate to Strategy Prompt from All Groups
(N = 53)*

Item	Pre-assessment	Post-assessment
62×23	43 (81%)	42 (79%)
$159 + 164$	51 (96%)	46 (87%)
$491 - 417$	41 (77%)	42 (79%)
$426 \div 7$	36 (68%)	40 (75%)

Thus, of the 53 participating students, between 2 and 19 students did not respond to the prompts for descriptions of their strategies for any given problem. Problems in Table 10 indicate this erratic response rate. The multiplication problem was administered first, followed by the

Instrument for Dissertation Study – Estimation Questions

addition, subtraction and division problems respectively. The uneven participation is more likely an artifact of problem type than of an order effect.

Using the algorithm described above, student responses were coded for complexity.

Simplicity of Computational Estimates

Table 11

Descriptive Statistics for the Complexity of Computational Estimation (CE) and Volume Estimation (VE) Problems

		Mean	Standard Deviation	Skewness	Kurtosis	Shapiro- Wilk's test (<i>p</i>)
PRE	CE Problems	3.04	1.23	0.88	-0.10	<.01
	VE Problems	3.37	0.75	1.98	3.82	<.01
POST	CE Problems	2.88	1.09	0.91	-0.01	<.01
	VE Problems	3.46	0.84	2.19	4.87	<.01

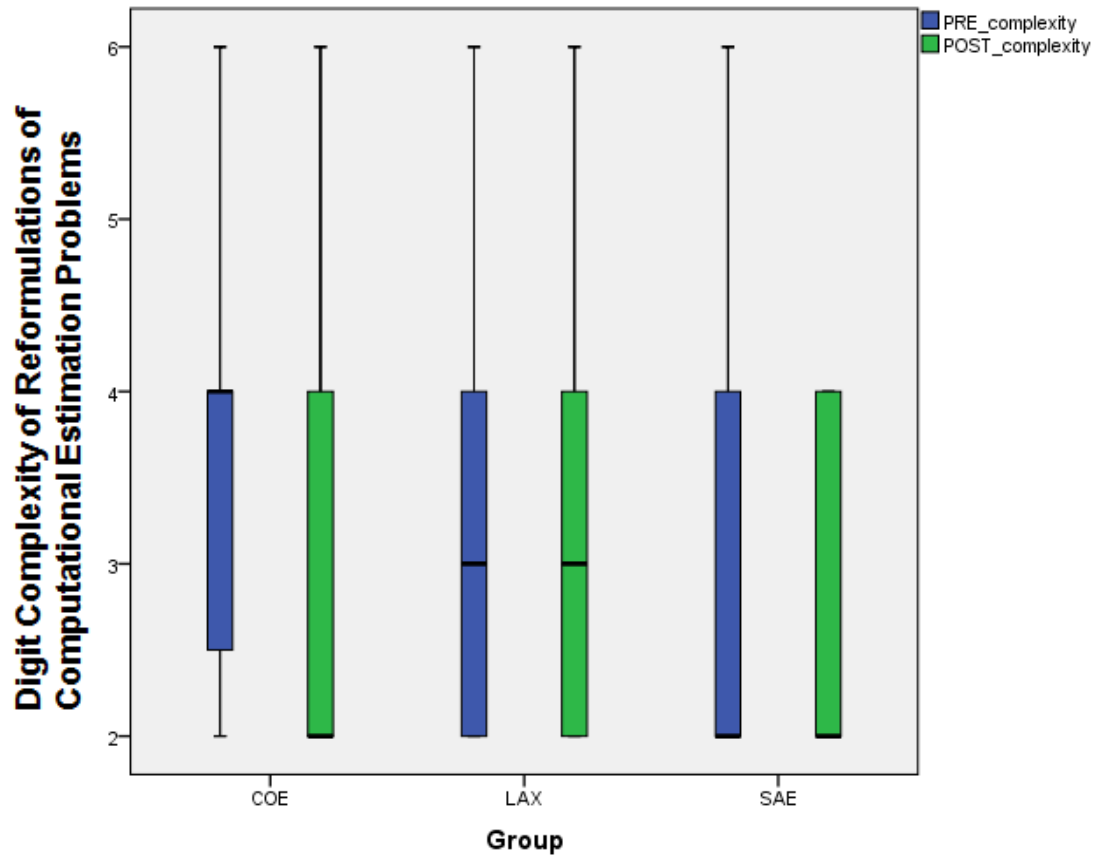


Figure 15: Boxplot of the Complexity of students' Computational Estimates

Table 12

Medians and Inter-quartile Ranges for the Complexity of Computational Estimates

	PRE		POST	
	Median	IQR	Median	IQR
COE	4	1.8	2	2
LAX	3	2	3	2
SAE	2	2	2	2

These values also failed Shapiro-Wilk's test for normal distribution, and were also transformed using Van der Waerden's formula. It should be noted here that values were not averaged within students before performing the simplicity analysis, as they were for the reasonableness analysis

Instrument for Dissertation Study – Estimation Questions

described in the previous section. This was because variation between complexity scores was highly sensitive to problem type.

Table 13

ANOVA for the Complexity of Computational Estimates, after Van der Waerden transformation

	<i>df</i>	<i>MS</i>	<i>F</i>
Group	2	3.02	3.54*
Time	1	0.27	0.85
Group × Time	2	1.20	3.87*
Between Group Error	123	0.85	
Within / Time Error	123	0.31	

Parentheses indicate Mean Squared Error.

Error was calculated across problems.

* indicates significant at the .05 level

A measure of the effect size of this relationship was obtained for the interaction (*partial* $\omega^2 = .03$). Simple main effects were calculated for time, per each group. Differences from pre- to post- were non-significant for the both estimation groups, including the COE group ($M = 0.39$, $SD = 1.06$), $t(23) = 1.81$, $p = .08$, $CI_{.95} -0.06, 0.84$, and the SAE group ($M = -0.05$, $SD = 0.95$), $t(42) = -0.36$, $p = 0.72$, $CI_{.95} -0.35, 0.24$. However, it was significant for the exact calculation group ($M = -0.13$, $SD = 0.46$), $t(58) = -2.18$, $p = .03$, $CI_{.95} -0.25, -0.01$. This is a surprising result, given the fact that the data for this group did not move (see Figure 15) but can be attributed to the ranking step involved in the Van der Waerden transformation. The movement *with respect to* the movement of the COE group is salient development.

Instrument for Dissertation Study – Estimation Questions

A more easily interpreted analysis of the data, shown in the following table, shows the mean complexity by problem for the computational estimation problems. For the SAE and COE groups combined, student descriptions of their strategies decreased on average:

Table 14

Pre- to Post-Assessment Complexity Scores for Both Estimation Groups (COE & SAE)
(*N* = 29)

Item	Pre-Assessment		Post-Assessment	
	Response Rate	Mean Complexity	Response Rate	Mean Complexity
62×23	23 (79%)	2.35	22 (76%)	2.09
$159 + 164$	27 (93%)	4.22	24 (83%)	3.29
$491 - 417$	22 (76%)	3.59	22 (76%)	3.27
$426 \div 7$	17 (59%)	2.76	20 (69%)	2.50

A close analysis of the digits themselves leads to some conclusions about how students are forming estimates. For example, the initial mean complexity score of 4.22 for the problem $159 + 164$ on the pre-assessment indicates that students were retaining more than four digits, on average. One common reformulation for this problem was $160 + 160$, which is one that many adults might choose (complexity = 4). However, eight students from these groups kept all of the digits in the problem, reformulating the problem with itself (complexity = 6). By the end of the unit, no students in either of the estimation groups used this exact-calculation strategy. Five of these eight students who originally kept all of the digits responded during the post-assessment, generating the following simpler reformulations: 200×200 ; 2×2 ; and 16×16 (complexity = 2; 2; and 4, respectively).

Table 15

*Overall Response Rate to Strategy Prompt from All Groups
(N = 53)*

Item	Pre-assessment	Post-assessment
Cube	38 (72%)	37 (70%)
Rectangular Prism 1	33 (62%)	39 (74%)
Rectangular Prism 2	23 (43%)	34 (64%)
Cylinder	13 (25%)	27 (51%)
Elliptic Cylinder	15 (28%)	19 (36%)
Rectangular Prism 3	11 (21%)	21 (40%)
Composite Shape	9 (17%)	12 (23%)

Items are listed here in the order that they were presented to the students. The third rectangular prism was designed as a recovery problem (Olshen, 1975, as cited in Reys et al, 1980). In other words, it was designed to be easier than the preceding problems in order to instill a sense of confidence in students if they felt challenged by the previous problems. However, the low rate both during pre- and post- assessment administrations may indicate the effect of students becoming fatigued.

Simplicity of Volume Estimates

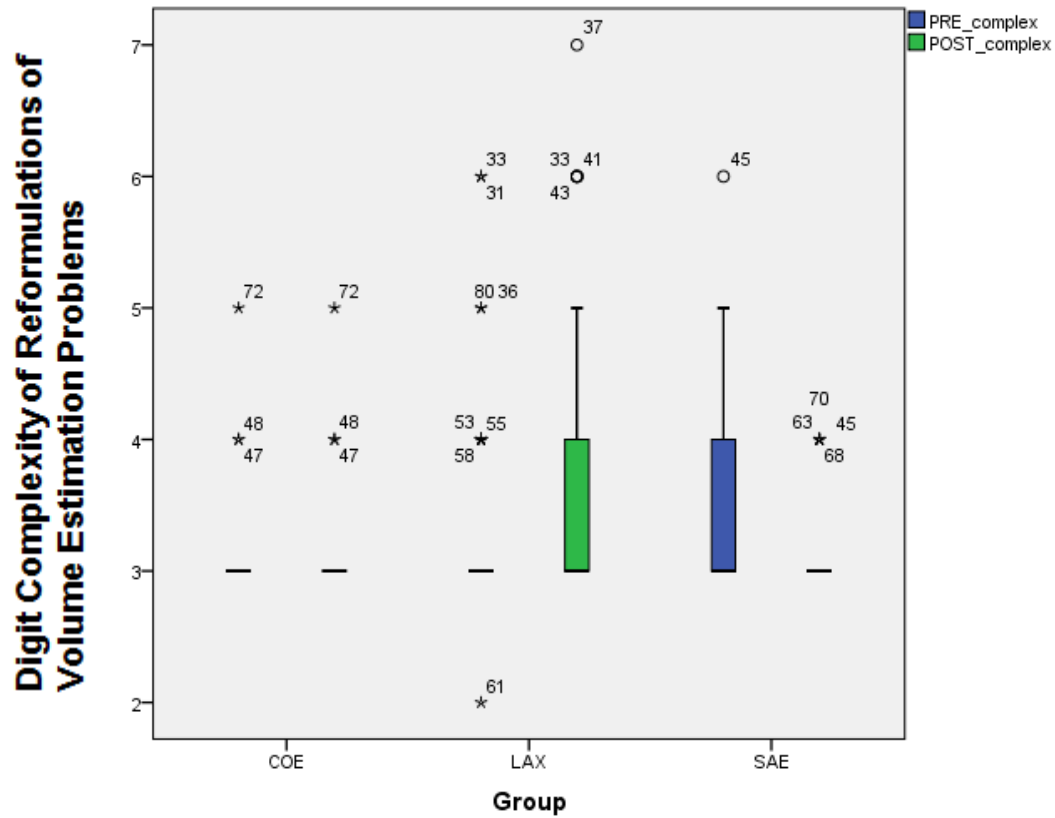


Figure 16: Boxplot of the Complexity of students' Volume Estimates

Table 16

Medians and Inter-quartile Ranges for the Complexity of Volume Estimates

	PRE		POST	
	Median	IQR	Median	IQR
COE	3	1	3	1
LAX	3	0	3	1
SAE	3	1	3	0

Table 17

ANOVA for the Complexity of Volume Estimates, after Van der Waerden transformation

	<i>df</i>	<i>MS</i>	<i>F</i>
Group	2	0.23	0.31
Time	1	0.20	0.63
Group \times Time	2	1.16	3.66*
Between Group Error	90	0.74	
Within / Time Error	90	0.32	

Parentheses indicate Mean Squared Error.

Error was calculated across problems.

* indicates significant at the .05 level

The group \times time interaction here seems to be an artifact of the SAE group's marginally simpler volume estimates after the intervention, and the LAX group's marginally more complex volume estimates after the intervention. A measure of the effect size of this relationship for the interaction was also obtained (*partial* $\omega^2 = .04$). Inspection of the boxplot and table above reveals that this corresponds to just one digit's worth of complexity in the difference between inter-quartile range between these two groups from pre-assessment to post-assessment. All medians for all groups were 3 for this kind of problem, both before and after the intervention. There were a large number of rectangles whose volume was to be estimated, and the frequent complexity score of 3 is a result of the simplification, to one significant digit, of the length, width, and height of each rectangle. As described further below, the SAE group tended to make slightly simpler estimates, for certain problem types, after the intervention. They used front-end estimation techniques slightly more frequently afterward.

Table 18

Instrument for Dissertation Study – Estimation Questions

Pre- to Post-Assessment Complexity Scores for Both Estimation Groups (COE & SAE)
($N = 29$)

Item	Pre-Assessment		Post-Assessment	
	Response Rate	Mean Complexity	Response Rate	Mean Complexity
Cube	18 (62%)	3.00	19 (66%)	3.00
Rectangular Prism 1	16 (55%)	3.50	18 (62%)	3.44
Rectangular Prism 2	11 (38%)	3.55	15 (52%)	3.53
Cylinder	4 (14%)	3.00	12 (41%)	3.25
Elliptic Cylinder	5 (17%)	3.60	7 (24%)	3.43
Rectangular Prism 3	2 (7%)	3.00	8 (28%)	3.00
Composite Shape	3 (10%)	3.00	6 (21%)	3.67

Visual inspection of this data reveals that there is no significant change from pre- to post-assessment in the mean complexity of student responses to the geometry items. A Wilcoxon Signed-Rank test confirms this ($Mdn_{pre} = 3.00$, $Mdn_{post} = 3.43$), $z = 0.41$, $p = .69$. The only notable change from pre- to post-assessment in the geometry items for these two groups is that three times as many students (12) responded to the cylinder estimation problem afterward as responded beforehand (4). The responses brought up the mean complexity of reformulations from 3.00 to 3.25. Of the four students who responded on the pre-assessment, only one included any estimate for π in his reformulation of the problem. He used $3 \times 10 \times 10 \times 20$ (complexity = 4) for a cylinder with a radius of 10mm and a height of 20mm. This reformulation produces an estimate for the volume of the sphere that has less than -5% error. On the post-assessment, students make similar reformulations, and 11 of the 12 respondents include an estimate for π . In fact, they all use 3 for π , which the prompt for the problem suggested that they do.

Nuance

As with simplicity, nuance was measured in terms of digits. The nuance score was a measure of how far the students' reformulations differed from algorithmic rounding. The initial conception of this algorithm had been to incorporate measures based on digits in the final estimates as well as digits in the reformulations. However, there was no systematic method for distinguishing whether an estimate that differed greatly from an algorithmic solution differed intentionally or resulted from computational errors. Thus, students' final estimates were excluded from this analysis, and nuance scores were obtained using only student reformulations, as with the simplicity (complexity) scores.

This analysis concerns only the computational estimation problems, since an analysis of volume estimation was deemed intractable. The following table shows the solutions that were determined to be "standard" solutions taught to students for estimation, based on left-most digit rounding.

Table 19

Algorithmic solutions to the Computational Estimation problems

Item	Reformulation	Solution
62×23	60×20	1200
$159 + 164$	$200 + 200$	400
$491 - 417$	$500 - 400$	100
$426 \div 7$	$400 \div 10$	40

Instrument for Dissertation Study – Estimation Questions

These reformulations are the closest to the algorithmic rounding solutions taught in schools, as described by Reys et al. (1982). This strategy matches the Truncate-SND technique. The choice to keep only the left digit in each problem in these reformulations is for simplicity. The following table describes the technique of analysis used to assign nuance scores to students' reformulation.

Table 20

A description of the Nuance scoring system for reformulations of Computational Estimation problems

Item	Reformulation	Nuance Score	Explanation
62×23	62×23	0	This is exact answer calculation.
62×23	60×20	0	This matches the algorithm.
62×23	60×25	1	The 5 is an adjustment.
62×23	6×2	2	The least-significant digits were ignored.

In order to generate the nuance scores, any reformulations that matched the original problem were scored as zero. This was considered exact answer calculation. As described earlier, an exact answer can be a valid estimate if it can be calculated mentally. However, the problems here were chosen to require sufficient digit manipulation so as to exceed working memory capacity (Cowan, 2001). Next, any problems matching the algorithmic rounding solution in Table 19, such as the reformulation of 60×20 for 62×23 were also given a score of zero. Finally, any digits that differed between the algorithmic rounding solution and the given reformulation were tallied for the nuance score. This included adjustments such as rounding a

Instrument for Dissertation Study – Estimation Questions

number to a more compatible solution. For example, reformulating the problem $426 \div 7$ to $420 \div 7$ would receive a nuance score of 3, because $420 \div 7$ differs from $400 \div 10$ in three digits. It represents a "compatible numbers" strategy (Reys et al., 1980), and can be considered a more nuanced solution than $400 \div 10$. The choice of digits might indicate an understanding that 42 is evenly divisible by 7. In sum, the nuance strategy was designed to tell whether students were finding a balance between retaining all of the digits and retaining the bare minimum number of digits when estimating.

Of the student responses to the items, 15 responses whose operator differed from the original problem. For example, on student reformulated $159 + 164$ as 160×160 . These responses were eliminated from the nuance analysis.

The following table shows the responses, before and after the intervention.

Table 21

Pre- to Post-Assessment Nuance Scores for Both Estimation Groups (COE & SAE)

Item	Pre-Assessment	Post-Assessment
62×23	1.04	0.64
$159 + 164$	2.67	3.29
$491 - 417$	1.70	1.43
$426 \div 7$	1.31	0.85

By this measure, the nuance of student estimates decreased for all operations except addition. Because this measure is intended to capture a shift away from very simple estimates, and also a shift away from overly detailed estimates (or exact computation). The following box plot gives a better picture of the nuance scores of students' estimates.

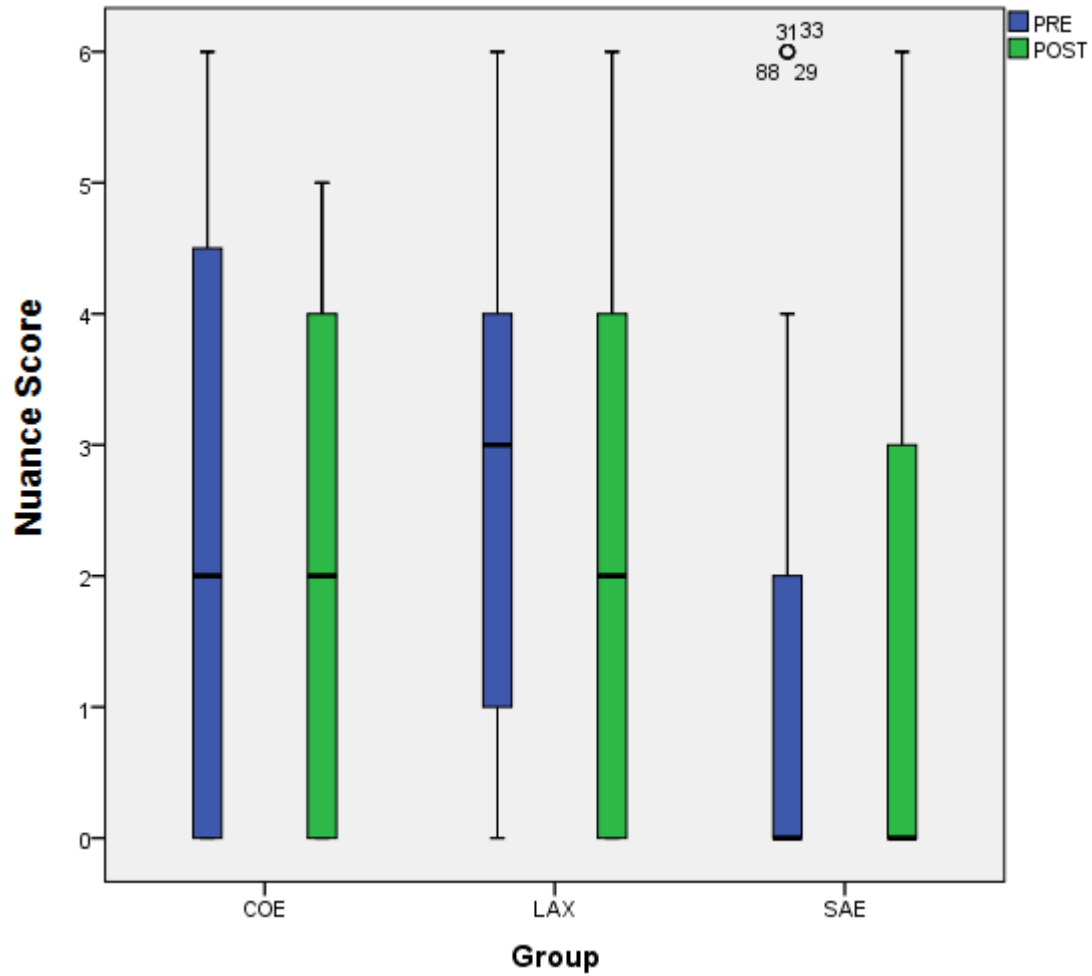


Figure 17: Boxplot of the Nuance of students' Computational Estimates

Table 22

Medians and Inter-quartile Ranges for the Nuance of Computational Estimation Strategies

	PRE		POST	
	Median	IQR	Median	IQR
COE	2	4.8	2	4
SAE	0	2	0	3
LAX	3	3	2	4

Instrument for Dissertation Study – Estimation Questions

Data shown in the box plot tell a story that fits with the reasonableness and complexity data. The COE students generate relatively nuanced reformulations, although some of these reformulations do not make sense mathematically, such as a reformulation of $491+417$ to $49+17^4$. Estimates of the LAX group are generally nuanced, while the SAE group's estimates exhibit no nuance whatsoever. The SAE group showed a tendency to seek simple, efficient answers, and their median nuance score of zero is likely a result of students seeking the simplest possible reformulation, which is often algorithmic rounding (nuance = 0).

Table 23

ANOVA for the Nuance of Students' Reformulations, after Van der Waerden transformation

	<i>df</i>	<i>MS</i>	<i>F</i>
Group	2	4.50	4.04*
Time	1	<.01	0.02
Group \times Time	2	0.23	0.89
Between Group Error	123	1.12	
Within / Time Error	123	0.27	

Parentheses indicate Mean Squared Error.

Error was calculated across problems.

* indicates significant at the .05 level

A measure of the effect size of the main effect for group was obtained (*partial* $\omega^2 = .03$).

A follow-up test with a Bonferroni-correction indicated that there was a significant difference in

⁴ This was a COE student's response to one of the extra 7 problems at the end of the assessment instrument.

means between the LAX group ($M_{post} = 0.18$, $SD = 0.87$) and the SAE group ($M_{post} = -0.15$, $SD = 0.81$), $p = .02$.

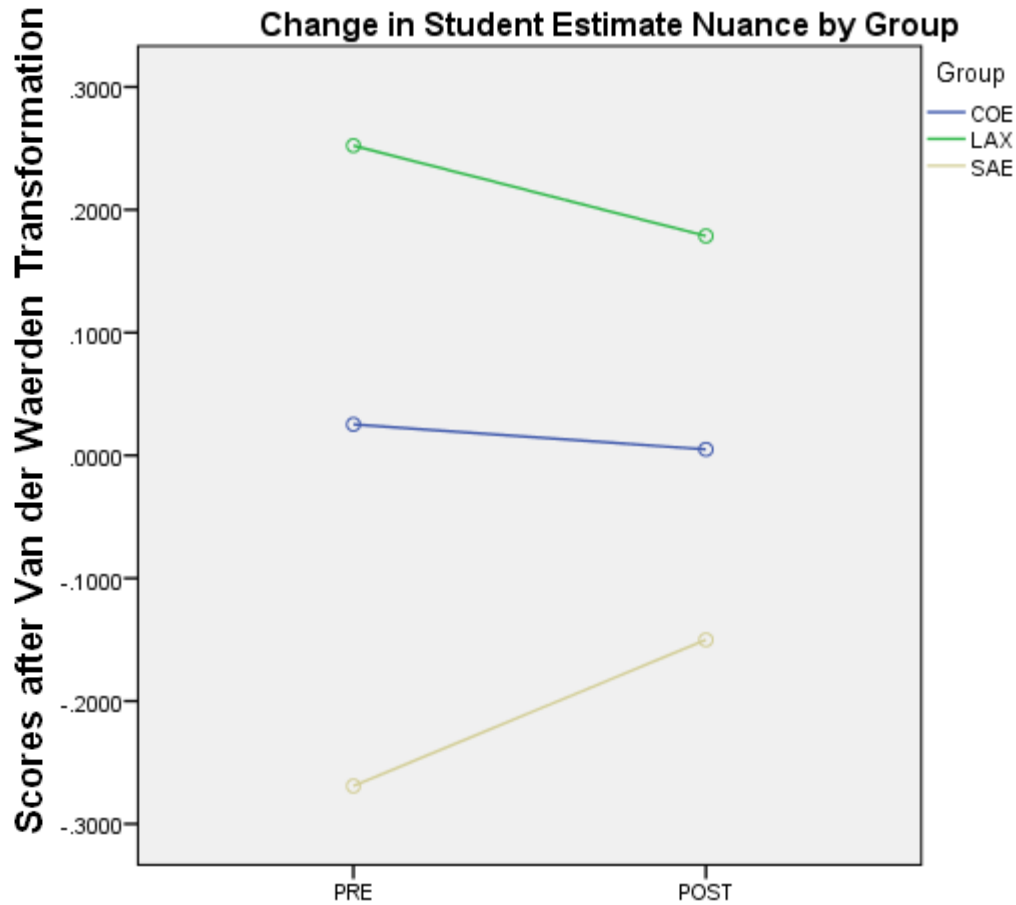


Figure 18: Pre- and post- scores for the three groups of students, shown as transformed data.

Evidence suggests that students did not use compensation strategies frequently if at all. In the section below entitled Adjustments to Estimates, there is a discussion of the adjustment feature in the software, which was designed to elicit student use of compensation.

The mid-range nuance scores for the COE group may be deceptive. Reformulations that do not make sense, such as $26 \div 7$ for the problem $426 \div 7$, received high scores (nuance = 5). This reformulation was made by one of the COE students during the pre-assessment. Her final

Instrument for Dissertation Study – Estimation Questions

estimate for this item was 25, and she did not respond to the same problem at all on the post-assessment.

As part of the nuance analysis, all of the students' reformulations based on exact computation were tallied. For example, if a student reformulated the problem 62×23 as 62×23 , this was considered exact computation. Occurrences of this went up slightly (*pre-assessment* 0 (0%)⁵; *post-assessment* 2 (2%)) for the LAX group, and down slightly for the SAE group (*pre-assessment* 4 (9%); *post-assessment* 2 (4%)). However, for the COE group, there was a large shift (*pre-assessment* 16 (37%); *post-assessment* 3 (8%)), suggesting that the students in this group lacked understanding at the time of the pre-assessment. The COE students may have misunderstood what they were being asked to do, or they may have not fully understood the meaning of the term estimate.

Simplicity and Nuance for the Exact Calculation Group

Since there was no main effect for time on the geometry problems in the omnibus test, there cannot have been a significant change in the reasonableness of student estimates for the volume or computation problems for the exact computation group.

⁵ Percentage based on response s for that group for that assessment session.

Simplicity**Table 24**

Pre- to Post-Assessment Complexity Scores for the Exact Calculation Group (LAX)
(N = 24)

Item	Pre-Assessment		Post-Assessment	
	Response Rate	Mean Complexity	Response Rate	Mean Complexity
62×23	20 (83%)	2.00	20 (83%)	2.10
$159 + 164$	24 (100%)	3.54	22 (92%)	3.86
$491 - 417$	19 (79%)	3.36	20 (83%)	3.15
$426 \div 7$	19 (79%)	2.57	20 (83%)	2.70

The above table indicates mixed changes in student responses. A Wilcoxon Signed-Rank test indicates no significant change due to the treatment ($Mdn_{pre} = 2.97$, $Mdn_{post} = 2.93$), $z = 0.73$, $p = .47$.

Instrument for Dissertation Study – Estimation Questions

Table 25

Pre- to Post-Assessment Complexity Scores for the Exact Calculation Group (LAX)
($N = 24$)

Item	Pre-Assessment		Post-Assessment	
	Response Rate	Mean Complexity	Response Rate	Mean Complexity
Cube	20 (83%)	3.00	18 (75%)	3.00
Rectangular Prism 1	17 (71%)	3.29	21 (88%)	3.38
Rectangular Prism 2	12 (50%)	3.17	19 (79%)	3.32
Cylinder	9 (38%)	3.78	15 (63%)	3.67
Elliptic Cylinder	10 (42%)	3.70	12 (50%)	4.08
Rectangular Prism 3	9 (38%)	3.00	13 (54%)	3.69
Composite Shape	6 (25%)	3.67	6 (25%)	4.83

Wilcoxon Signed-Rank analysis suggests that what appears to be a slight increase in complexity of student responses is non-significant ($Mdn_{pre} = 3.29$, $Mdn_{post} = 3.67$), $z = 1.78$, $p = .08$.

Nuance

Table 26

Pre- to Post-Assessment Nuance Scores for the Exact Calculation Group (LAX)

Item	Pre-Assessment	Post-Assessment
62×23	1.50	0.85
$159 + 164$	4.62	4.10
$491 - 417$	3.21	2.06
$426 \div 7$	1.58	1.32

As with the estimation groups, the amount of nuance in student estimates has decreased. This is not significant ($Mdn_{pre} = 2.39$, $Mdn_{post} = 1.69$), $z = -1.83$, $p = .07$, according to a Wilcoxon Signed-Rank test.

Adjustment to Estimates

A more direct measure of nuance in estimation strategies is the students' use of the adjustment. The interface for the assessment instrument featured a check box with the following caption: *Check this box if you want to adjust your estimate.* If the user clicked this box, a text-box would appear that would allow the user to enter text describing the strategy he or she used to make the adjustment. The LAX group disproportionately used this feature, with 25 uses on the pre-assessment and 18 uses on the post-assessment. Meanwhile, the other two groups used it only 9 times during the pre-assessment and not at all on the post-assessment. All of this usage by the LAX group was by just 5 of the students.

The adjustment feature was used almost exclusively for place value adjustment. Of the 52 times that it was used, 50 usages described appending, or "adding" zeroes. The other two usages are unclear. One of the COE students typed "I estimated 126 by getting 120" as a description of her adjustment and she gave 120 as her final estimate. One of the LAX students wrote "SIMPLIFY FINAL ANSWER" in all capital letters for the problem $159 + 164$ on the post-assessment. This student's reported strategy was to reformulate the problem as $160 + 160$. She reports a final estimate of 400. Thus both of these responses seem to be form of pseudo-estimation, or rounding the final estimate simply to make it a round number.

Theme: Estimating the Volume of a Cylinder

Analysis of the pre- to post-assessment show a change in the way students estimate the volume of a cylinder. Among the estimation groups, students demonstrated more use of π during the post-assessment than during the pre-assessment.

Eight students from the COE and SAE groups responded to the cylinder volume estimation question on the pre-assessment. Only one made an estimate that was within a $\pm 15\%$ tolerance: an estimate of $6,000 \text{ mm}^3$ for a shape whose exact volume was 6283.19 mm^3 . The remaining estimates, in decreasing order, were $2,000 \text{ mm}^3$, $2,000 \text{ mm}^3$, $1,000 \text{ mm}^3$, 200 mm^3 , 60 mm^3 , 30 mm^3 , and 20 mm^3 . Clearly, a mixture of place value and front-end digit difficulties are responsible for all of these underestimates. Half of these students explained their strategies, and only the student who made the reasonable estimate included an estimate for π in the reformulation he described. He used 3, as the prompt for this problem suggests.

By contrast, nine of the students in the LAX group described their estimation strategies during the pre-assessment and every single description included an estimate for π . This does not generalize to the entire group; one student in the LAX group saw the teacher write π on the board a couple of days later and asked, "What is that wavy thing with legs?" Nevertheless, a significant fraction of the LAX students knew how to operate with π before the intervention.

On the post-assessment, 17 students from the COE and SAE groups responded to this problem, 12 of the responses included mathematical descriptions of their strategies, and 11 of these included an estimate for π . There were also five estimates of $6,000 \text{ mm}^3$, which constituted the only reasonable estimates according to the Estimation Calculator default criterion of a $\pm 15\%$ tolerance. One student made an estimate of $5,000 \text{ mm}^3$, which might be considered reasonable by more liberal definition.

Instrument for Dissertation Study – Estimation Questions

To summarize, among students in the estimation groups, the ratio of estimates that were reasonable rose from 1:8 during the pre-assessment to 5:17 during the post-. This corresponded to a rise from 1:4 descriptions that included π to descriptions overall in pre-assessment to a ratio of 11:12 during the post-. The students' broad acknowledgement of π in the post-assessment clearly did not guarantee their formulation of reasonable estimates. Their estimates were still distorted by place-value and confusion over squaring the radius. However, only one student (representing 13% of the respondents) in the pre-assessment used π , and used it correctly, and 5 students (29%) used π correctly in the post-assessment. This suggests that students established how to solve problems using π .

Results of Student Interviews

Six students were chosen at random to be interviewed before and after the unit. The pre-interview consisted of 6 questions, including 2 volume estimation questions. For certain problems, including the second question (896×19), students were given the chance to use the Estimation Calculator on a laptop to receive interactive feedback on their estimates. The post-interview consisted of the same first three questions, and the same last two questions. However, a new question was added in the middle of the interview to assess whether students could now estimate the time needed to print a small pencil holder in plastic. This is referred to as the pencil holder problem in the following section. The prompt for this question (see Appendix B) hinted at two possible strategies for making this estimate. On the one hand, students could use a multi-step mathematical strategy with the information given. On the other hand, they could simply use the pencil holder itself, along with the set of mental anchors developed during the unit to estimate time-needed-to-print. Students used a mixture of these two strategies.

Instrument for Dissertation Study – Estimation Questions

The following section summarizes the student interviews, including charts showing the error of student estimates, log-transformed with the formula at the beginning of this chapter. At the beginning of each of the following summaries is the student's Mean Absolute Magnitude Error (MAME), which is based on this same transformation. Thus, a MAME of 1 indicates that on average, a student made estimates that were one order-of-magnitude from the actual answer. These range from 0.19 to 1.09.

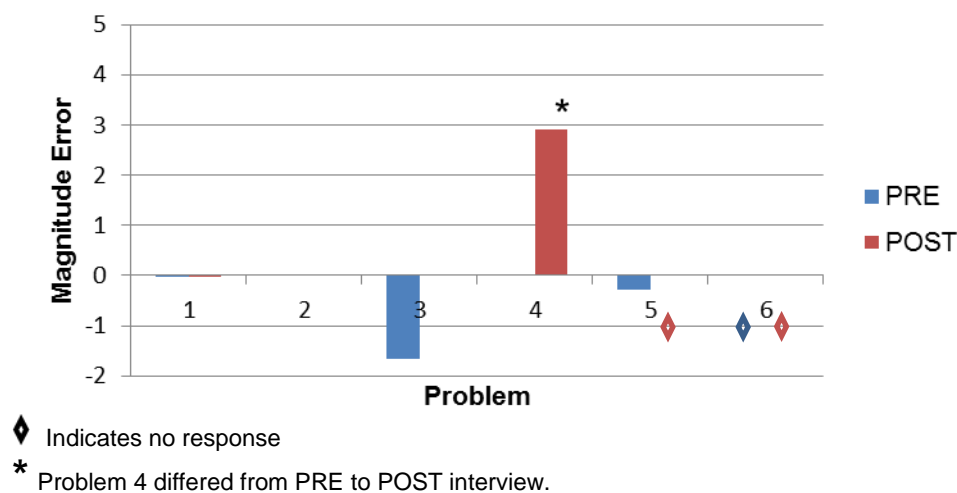


Figure 19: Estimate Errors for student COE0401

The student with id COE0401 answered 9 of the 12 questions, with a MAME of 0.56. She was thorough and systematic, and she worked through problems slowly. The diamonds indicating unanswered problems are all at the right-hand side of her chart, corresponding to items that she did not have time to complete.

She reformulated the first problem as $400 + 80$ for an estimate of 480 during both the pre- and post-interview. For problem 5 on the pre-interview she was given the chance to use the Estimation Calculator to check her estimate. Her original estimate was actually an exact calculation of the answer, using a traditional stacked multiplication strategy, although she made a

Instrument for Dissertation Study – Estimation Questions

computational error. The Estimation Calculator showed her that this estimate (104) was too low. The problem (shown in Appendix B) required estimating the volume of a rectangular prism whose dimensions were $17\text{mm} \times 4\text{mm} \times 3\text{mm}$. The actual volume is 204mm^3 . After seeing the error feedback bar, she started over, using all the digits in the original problem to arrive at the exact answer. The following dialogue occurred after she saw the output on the Estimation Calculator.

COE0401: So should I redo it or something?

Interviewer: If you want to redo it, there's a button there that says 'Try Again.' So whatever strategy you think would be best to try again.

COE0401: So, I guess I'll just try doing the math again.

She wrote addition and multiplication problems in stacked column format.

Her error on the multi-step, pencil-holder problem in the post interview involved not understanding how to divide by 1000. This resulted in an estimate that was much too large (50005 minutes). She then made a completely separate estimate of the amount of time needed to print the object based on looking at the object itself, writing "1 hour, 30 min". The exact, computed answer is 61.55 minutes. This problem took up a significant amount of the interview, and she did not have time to provide estimates to the remaining two problems. This means she did not have time to revisit the $17\text{mm} \times 4\text{mm} \times 3\text{mm}$ problem, described above.

At first, this student chose reformulations that hardly differed from the original problems.

During the pre-interview, she depended heavily on stacked, columnar multiplication strategies, such as on the second problem, where she reformulated 896×19 as 900×19 , and then computed an exact answer on paper. During the post-interview, she made more frequent use of front-end strategies and retained fewer digits. This time, she reformulated 896×19 as 900×20

Instrument for Dissertation Study – Estimation Questions

which she then solved in her head. This is a subtle shift, but it is also evident in her solution to the pencil holder problem, and may indicate a preference for simpler reformulations after the intervention.

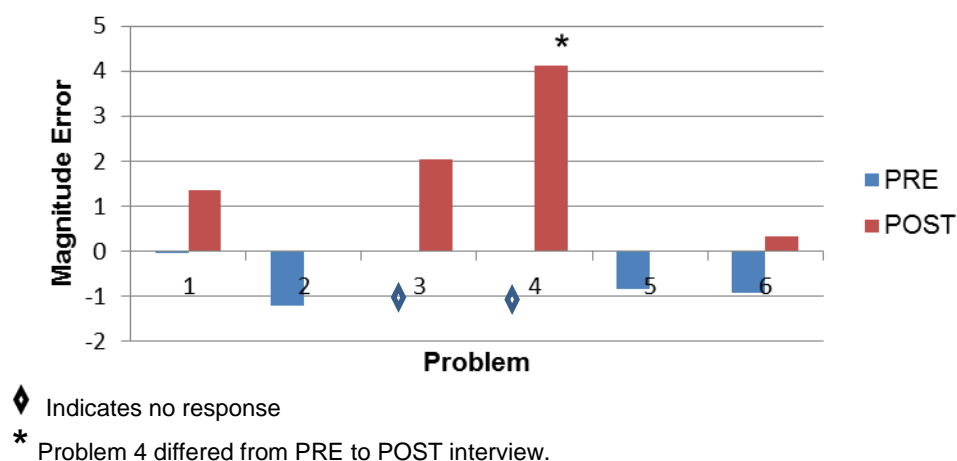


Figure 20: Estimate Errors for student COE0402

Unlike COE0401, student COE0402 worked quickly. As the chart indicates, this student struggled significantly with place value, and almost every answer she provided included an order-of-magnitude error. She had the highest overall MAME (1.09) of any of the interviewees. She left only two questions unanswered (3 and 4). Both of these were problems she encountered and chose to skip during the pre-interview. Both problems involved calculation with percentages. Question 3 on the post-interview was identical to question 3 the pre-interview, and she did answer it during the post. Her estimate was based on a valid front-end strategy, with a two-digit error in place value, indicating she did not understand the meaning of the % sign. Percentages were not directly taught during the intervention, and so there is no reason to expect improved performance. Her willingness to answer it in the post-interview may indicate an increased level of self-efficacy with estimation (Bandura, 1977) or greater personal comfort responding to the interviewer, who had spent the previous week in her mathematics class.

Instrument for Dissertation Study – Estimation Questions

During the pre-interview, student COE0402 was already familiar with an algorithmic approach to rounding for estimation. She underlined left-most digits and wrote zeroes on top of the other digits, even for the number 1000. For 1000, she underlined the 1, and then wrote a 0 on top of the 0 that already existed in the hundreds place.

She used the Estimation Calculator heavily for feedback. When asked to make an estimate for 896×19 , she reformulated the problem as 900×20 and then said "maybe, like, a thousand something." She then typed the following sequence of estimates into the Estimation Calculator: { 1056, 1098, 2000, 2098, 3000, 3098, 4000, 4098, 5000, 9000, 11000, 2200 } and then said, "I don't know." For the same problem on the post-interview, she did not write anything on the paper except her final estimate. Nor did she make multiple attempts at estimates with the Estimation Calculator. Her first and only estimate was 18000. This was a reasonable estimate; the actual answer is 17024.

As with the other COE student, she had difficulty with the pencil holder problem on the step requiring division by 1000. She simply skipped this step. Her final estimate was four orders of magnitude too large (800000 minutes). She had trouble distinguishing between operations, including determining the order of operations or PEMDAS. Her estimate seems to resulted from trying to subtract 3600 from 5200. She used a front-end (addition) strategy and then a combining-zeroes (multiplication) strategy as a method for subtraction. After this, when asked how long she thought it must have taken to print the pencil holder, she responded "about an hour."

Instrument for Dissertation Study – Estimation Questions

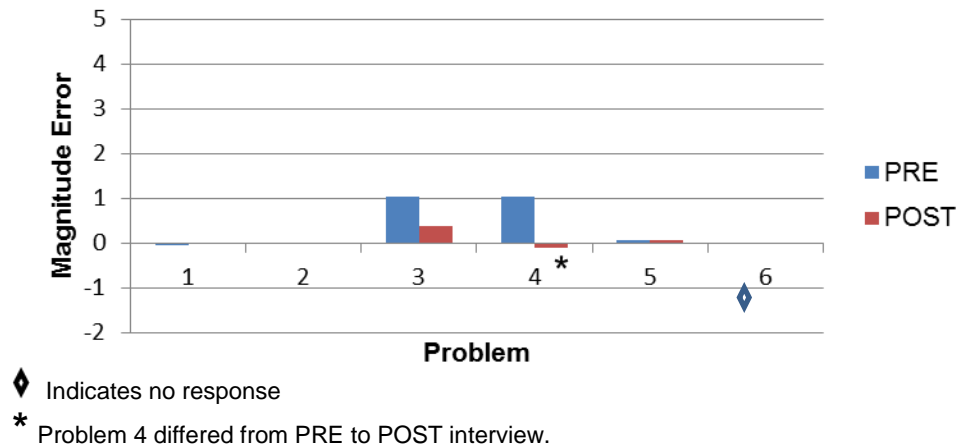


Figure 21: Estimate Errors for student SAE0317

Student SAE0317 was an efficient worker, with a well-developed number sense. She seemed to understand that the purpose of estimation is to simplify and streamline computation. She wrote fewer notes than any of the other interviewees, and except for two mistakes on the percentage problems (questions 3 and 4 on the pre-interview), all of her responses were within an order of magnitude of the exact answer (MAME 0.25). She did not complete the final problem during the pre-interview because her interview was cut short. However, she completed the entire post-interview in half of the allotted 15 minutes, due mostly to the fact that she solved the pencil holder problem without any computation.

Interviewer: It's an outer cylinder minus an inner cylinder. The outer cylinder has a radius of 20 millimeters and a height of 30 millimeters. The inner cylinder has a radius of 10 millimeters and a height of 30 millimeters.

SAE0317: I'm going to guess about 60-something minutes.

Interviewer: Ok.

SAE0317: No-no-no, 50-something minutes.

Instrument for Dissertation Study – Estimation Questions

Her response to the pencil holder question indicates her grasp of the purpose of estimation. The exact, computed answer is 61.55 minutes. Student SAE0317 used a valid reference-point method of solving the problem (Joram et al., 1998), which many of the other students used as a follow-up strategy. In other words, students attempted the multi-step computational method first and the use of the reference point second.

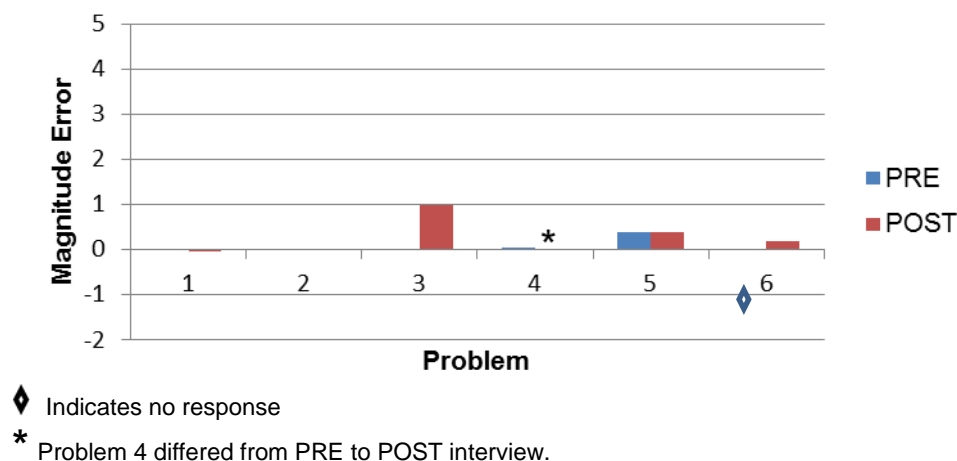


Figure 22: Estimate Errors for student SAE0306

Student SAE0306 made the most accurate estimates of any of the six interviewees (MAME 0.19), although his estimation strategy was not always efficient. There was some change in his approach between the pre-interview and the post-interview. During the pre-interview, he used lattice multiplication for all but the first problem. Lattice multiplication involves writing the same number of digits as traditional, stacked multiplication, as well as drawing a lattice, or grid. The lattice contains diagonal slashes dividing each cell into upper-left and lower-right halves to help manage the place value operations in substeps of the problem. In answering the same questions on the post-interview he relied much less heavily on this technique, instead using front-end multiplication strategies that produced an accurate number of zeroes. Because this front-end strategy was taught directly as part of the intervention, this

Instrument for Dissertation Study – Estimation Questions

change is explicable. Front-end division with large numbers was not taught as part of the unit. Nevertheless, student SAE0306 demonstrated an understanding of eliminating zeroes as a technique for dividing by a power of 10. He used lattice multiplication six times during the pre-interview. He fell back on lattice multiplication only three times during the post-interview questions, even though the pencil-holder problem added six additional multiplication operations (using the algorithm he chose) when compared with the same problem on the pre-interview.

This student was the only one to accurately divide by 1000 in the post-interview. He said, "Sixty thousand divided by one thousand. It's going to be ... sixty." There are at least two valid techniques that lead to this solution. One involves recognizing that $60,000 \div 1,000$ is analogous to $60x \div 1x$, where x is any value. The other involves eliminating an equal number of zeroes from the ends of both the divisor and the dividend and then manipulating the front-end digits as with any other division problem. This student may have used either of these techniques. What is evident is that his chosen strategy was efficient. He performed the computation in his head in approximately 2 seconds.

The fact that SAE0306 solved more problems during the post-interview than during the pre-interview suggests that he may have successfully incorporated efficient multiplication estimation strategies taught during the unit into his repertory of strategies. There is evidence that he knew a front-end multiplication strategy during the pre-test. By the end of the unit, possibly as a result of more instruction and further practice, these strategies were more central to his repertory.

Instrument for Dissertation Study – Estimation Questions

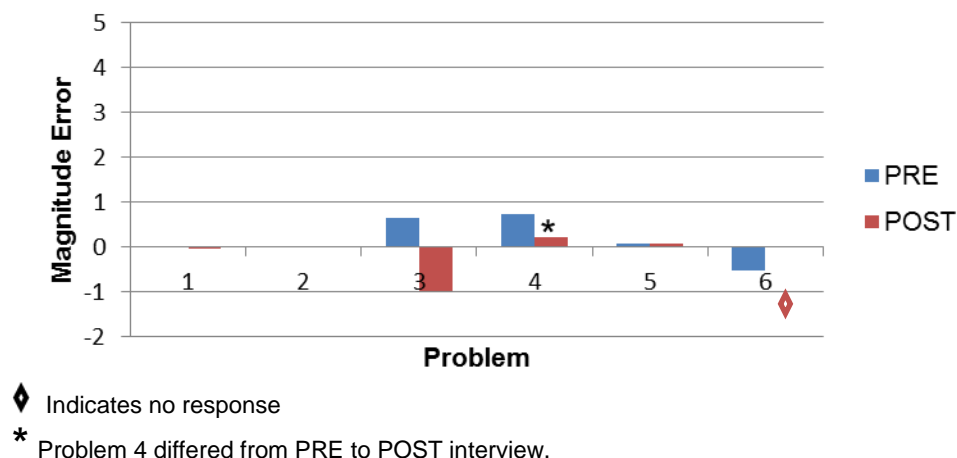


Figure 23: Estimate Errors for student LAX0702

This student already used effective front-end rounding strategies for multiplication and addition during the pre-interview. Her greatest difficulties occurred with problems involving percentages. Her MAME was 0.30, and she exemplifies the kind of student who did not need much help learning to make estimates, at least ones involving addition and multiplication. During the post-interview, she attempted to solve $108000 \div 1000$ using long division, resulting in a solution of 99.9. This solution has the right magnitude, but was clearly the result of flawed execution of the long division algorithm. When asked, she identified minutes as the proper units for the solution to the pencil holder problem and provided 99.9 minutes as her final estimate for the pencil holder problem (The exact solution is 61.55 minutes).

Instrument for Dissertation Study – Estimation Questions

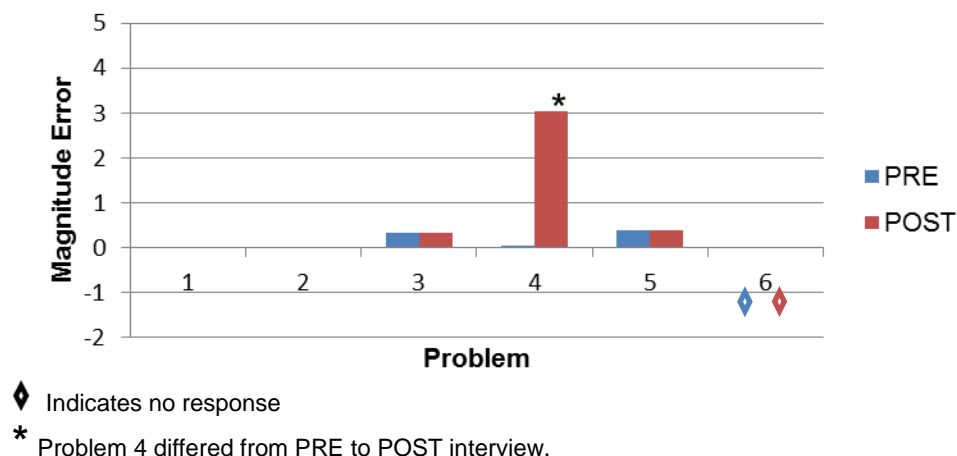


Figure 24: Estimate Errors for student LAX0725

This student was remarkably consistent. As with student LAX0702, he struggled with percentages. His MAME was 0.46, which was due mostly to his difficulties with place value on the pencil holder problem. On all of the remaining problems, his responses remained exactly the same from pre- to post-interview. This included his use of a front-end estimation strategy, which involved retaining one or two of the front-end digits and correctly manipulating the remaining zeroes.

He worked slowly, not reaching problem 6 in either the pre- or the post-interview. During the pencil holder problem, he attempted to use long division to solve $330,000 \div 1,000$. His erroneous result was 33,000. This division was part of solving a formula to convert the volume estimate to a time estimate. The units are intentionally not included in the written prompt for the pencil holder question (see Appendix B), and students were asked verbally if they remembered the units (minutes) from the corresponding classroom activity. The following dialogue occurred as student LAX0725 used 33,000 as an intermediate result for the formula $Time = 2 \times V/1000 + 5$.

Interviewer: So you're going with thirty-three thousand.

Instrument for Dissertation Study – Estimation Questions

LAX0725: And then I times it by two. ... Sixty-six thousand. And then, plus five would be sixty-six thousand and five.

Interviewer: Ok. Do you remember what the units were for time, when we were talking about that?

LAX0725: Millimeters.

Interviewer: For time?

LAX0725: For time? Oh. [long pause]

Interviewer: It's okay if you don't. I was just asking.

[LAX0725 shakes his head.]

The interviewer skipped the follow-up question at this point. All of the other interviewees were asked to make a ballpark estimate for time-needed-to-print based on simply looking at the pencil holder, if they had not already reached a reasonable estimate computationally. Therefore this student did not demonstrate whether his intuition and mental anchors from the previous week would lead him to a reasonable estimate. This is an unfortunate missing data point.

Clearly, he did not choose minutes as a unit to accompany his computed estimate of 66,005. It is impossible to know what he would have said if he had arrived at an estimate with a more reasonable magnitude.

Within Group Observations

The following table (Table 27) summarizes the observations of students within each group, under four themes that emerged from the analysis. The first two themes are mathematical processes that highlight differences between groups, within which students from each group are characterized. These themes are the understanding of place value, and division by 1000. The

Instrument for Dissertation Study – Estimation Questions

right-half of the table describes any changes in student strategies or descriptions of their strategies, both for computational and volume estimation.

Table 27

Observations within groups for interview estimates.

Group	Understanding of Place Value	Division by 1000	Change in CE strategies from PRE to POST	Change in VE strategies from PRE to POST
COE	Weak. Both students struggled with percentage problems and division by 1000.	Struggled. Both students substituted algorithms from other operations	Slight shift toward front-end estimation strategies (COE0401).	None.
SAE	Strong, with the exception of some percentage problems.	Efficient. Both students found efficient algorithms, with or without division.	Slight shift toward front-end estimation strategies (SAE0306).	None.
LAX	Mixed. Weak on percentage problems and division by 1000 (LAX0725).	Mixed. Both students attempted long division and miscalculated the quotient.	None.	None.

The standard, or COE group, generally lacked confidence in their abilities. The group included a number of students who confused the techniques needed to solve problems with the four basic operations. Student COE0402 was one of these students. The other student from the group, COE0401, had abilities that were roughly average compared to her peers. Her tendency to start over from the beginning on a difficult problem when she received feedback on its

Instrument for Dissertation Study – Estimation Questions

reasonableness suggests that she trusted the algorithm more than a sense of the reasonableness of the result. In this sense, COE0402 represented the opposite tendency, since she used the Estimation Calculator to repeatedly revise an answer simply based on feedback. These can be seen as two different responses to a lack of confidence in one's answer.

In a follow-up interview, the teacher acknowledged the difficulties of the COE group, particularly with addition estimation, saying, "Those standard kids needed a lot of practice with the addition estimation." This highlights the fact that these students often struggled with mathematical processes that are taught in upper elementary grades. The standard (COE) class was hampered in making sound volume estimates because of a failure to generate reasonable computational estimates.

The SAE and LAX groups were both advanced groups, and these students generally knew efficient ways of manipulating the numbers involved in estimation. The SAE group interviewees in particular had confidence in their own answers, and offered them quickly and confidently. The LAX group was more careful, particularly in solving the multi-step pencil holder problem. While the interviewees were not fully representative of their respective classroom groups, they each embodied key characteristics from the group at large.

Summary

It is possible to summarize the characteristics that emerged. Overall, the LAX group was characterized by *proficiency*. Interviewees in this group used proper mathematical techniques, such as long division, even if they made some errors in execution. Their estimates were fairly reasonable overall. The SAE group was characterized by *efficiency*, particularly after the intervention, including their use of rapid front-end estimation strategies and a print-time estimation strategy based on the use of mental reference points. Meanwhile, the COE group was

Instrument for Dissertation Study – Estimation Questions

characterized by *uncertainty*. This manifested as slow, careful redoing of basic problems from scratch when errors emerged (COE0401) and rapid, repeated trials that resembled floundering on the Estimation Calculator.

Across Group Observations

As the table above (Table 27) shows, no students demonstrated growth on volume estimation strategies. The volume estimation problems for the interviews were designed so that compensation strategies could be used to adjust digits to more closely approximate the exact answers. Although compensation was part of the week-long curriculum, students from the two estimation treatment groups did not adopt a compensation strategy when calculating volume. Neither did the LAX group. This would have indicated the understanding of applying computational estimation to volume estimation. The interviewees in the estimation treatment groups did, however, tend slightly more toward the use of front-end estimation strategies for the computational estimation problems during the post-interview than they did during the pre-interview.

Mean Absolute Magnitude Error for All Interviewees

The following table (Table 28) shows the mean absolute magnitude error (MAME) of the students' estimates during the interviews. These error scores are given both including and excluding the pencil holder problem for the post-interview, because estimates for this problem were several orders of magnitude away from the exact answer for three of the students. This large error was due to difficulty during the division by 1000 step, as discussed below.

Table 28

Instrument for Dissertation Study – Estimation Questions

Mean Absolute Magnitude Error (MAME) for student interview estimates.

Student ID	All Problems	PRE	POST	POST: Excluding Pencil Holder Problem
COE0401	0.56	0.41	0.75	0.03
COE0402	1.09	0.75	1.32	0.76
SAE0317	0.25	0.44	0.10	0.10
SAE0306	0.19	0.09	0.27	0.32
LAX0702	0.30	0.33	0.26	0.27
LAX0725	0.46	0.16	0.76	0.19

There are three themes that emerged from the interviews. These are the simplicity of the multiplication estimation strategies, students' struggles with division by 1000, and students' acceptance or rejection of their own printing-time estimate for the pencil-holder problem. The following subsections examine these themes in more detail.

Simpler Multiplication Estimation

Students in the estimation groups used simpler multiplication estimation strategies during the post-interview than during the pre-interview. Student SAE0306 relied heavily upon lattice multiplication initially and then used front-end strategies in the post-interview. His strategy use most clearly demonstrates a shift. However, COE0402's quick front-end strategy for multiplication on the post-test Estimation Calculator problem and COE0401's slight shift from stacked multiplication to front-end multiplication across all of the problems support this to a lesser extent. Meanwhile, little change in strategy appeared among the two students in the exact computation (LAX) group. Student LAX0702 already used front-end estimation productively

Instrument for Dissertation Study – Estimation Questions

during the pre-interview and did so as well during the post-, and LAX0725 used an equal mix of front-end and stacked multiplication on both pre- and post-interview.

Struggles with Division by 1000

On the pencil-holder question, division by 1000 was problematic for four of the six interviewees. Only one student (SAE0306) performed an efficient front-end division strategy, solving $60,000 \div 1000$ correctly without rewriting the problem. Another student (SAE0317) avoided the need to divide by 1000 by solving the pencil holder problem with a reference-point (Joram et al., 1998) volume-estimation strategy and avoided any computation. The remaining four students struggled with this problem. One substituted a subtraction algorithm (COE0401), one used a hybrid addition / multiplication algorithm (COE0402) and two attempted, and struggled with, long division (both LAX students).

Acceptance or Rejection of the Time Estimate

Five of the six students provided reasonable estimates, with the proper units, for time needed to print the demonstrated pencil holder in plastic. These estimates ranged from "50-something minutes" to 99.9 minutes. Each of these students gave their estimate in response to the question, "How long do you think it took to print this pencil holder?" The two COE students both abandoned their numerical estimates (50,005 and 80,000) at that moment and gave estimates based on what made sense to them ("1 hour and 30 minutes" and "1 hour"). The two SAE students came up with reasonable estimates, one through correctly applied volume and computational estimation strategies and the other through a reference-point strategy, in other words estimating based on prior knowledge ("50-something minutes"). One of the LAX students made computational errors and was not asked the follow-up question about estimated time after

Instrument for Dissertation Study – Estimation Questions

he had reached an unreasonably large estimate for time (66,005). Finally, the other LAX student made a division error that was small enough so as to leave her with a reasonable estimate for time (99.9). She reasonably concluded that this estimate could be interpreted as 99.9 minutes.

One student from each advanced group, LAX and SAE, made a fully computed, reasonable estimate. By contrast, both students in the non-advanced, COE group made estimates based on computations that were sufficiently flawed as to be unreasonable.

However, both students in the COE group revised their estimates so as to be reasonable. These revisions were based on a shift of focus from the multi-step problems on which they had been working to volume estimation based on a tangible artifact with which they had some experience. This suggests that exposing the COE-group students to 3D-printed artifacts, as well as demonstrating a 3D printer in action, and focusing on time needed to print these items did provide a foundation for reasonable estimation on a desktop-manufacturing related transfer task. Meanwhile, the students in the SAE and LAX groups who reached reasonable estimates computationally and accepted them also demonstrated this estimation ability. Since no data exists on what kinds of printing-time estimates these students would have made before the unit, repeated measures comparisons cannot be drawn.

Summary

In summary, salient results from the interview concerned multiplication, division, and a sense of reasonableness in the applied context of 3D printing. First, some students from among the estimation groups made simpler multiplication estimates during the post-interview than during the pre-interview. Second, there were almost universal struggles with division by 1000 across groups. Third, students were willing to accept an estimate for printing time if and only if it was reasonable.

Instrument for Dissertation Study – Estimation Questions

As with the results described in prior sections, these results tend to highlight the differences between advanced and non-advanced groups more than they do the difference between students who did and did not receive instruction on estimation.

Snapshots from the Intervention: Students' 3D Products

Six students were chosen at random to have their pencil holder designs 3D printed in plastic on the desktop manufacturing device. These students were chosen from among the non-interviewees to avoid any influence this might have on post-interview responses. Students' designs ranged in size from one that could comfortably hold two pencils (and fit three, but fell over when the third one was added) to a pencil holder that was too small to actually hold a pencil. It was discovered that this very smallest object could in fact hold a pen if the pen was put in cap-first. The object was just large enough to grip the cap of a standard ball-point pen, and if placed carefully on a desk, it could be balanced with the pen pointing straight up.

Students were constrained to designing pencil holders that could be printed in one class period. As one of the class activities leading up to the actual pencil holder design, students performed exercises using a conversion formula that related the volume of a pencil holder to its print time. This conversion activity was part of the curriculum for all three groups. It included a discussion of units, which were mm^3 for the volume of the pencil holders and minutes for the print time. This conversion was based the result of linear regression from five pencil holders that were printed before the unit began. It was an imperfect estimator ($r^2 = .82$) that was adjusted to make the numbers easier for the students to manipulate: $\text{Time} = 2 \times V/1000 + 5$. Students were told that their pencil holders should be less than $20,000 \text{ mm}^3$, in order to print in 45 minutes or less.

Instrument for Dissertation Study – Estimation Questions

The interventions were originally planned so that students would estimate (or calculate) the volume of their designs as they designed them. However, few students followed the directions on the worksheet that accompanied the design activity, which instructed them on how to estimate or calculate the volume. The result was that most students' designs violated the principles of informed design (Burghardt & Hacker, 2004).

To reinforce the processes of volume calculation and estimation, a worksheet was developed using the first two pencil holders that were printed and was given to students the day after the first pencil holders design day. One version of the worksheet included instructions to *estimate* the volume, while the other version stated *calculate*. The LAX and SAE groups completed the calculation and estimation versions, respectively, within 15 minutes at the start of class. The COE group required detailed instruction with the teacher, the collaborative teacher, and the researcher guiding clusters of five or six students through each step of the estimation process on the estimation version of the worksheet.

As a conclusion to the worksheet activity, the teacher and researcher led a class discussion about the volume of the pencil holders using base-10 blocks. A base-10 unit cube is 1 cm long on each edge, and so provides a useful reference point for a volume of $1,000 \text{ mm}^3$. The counter-intuitive fact that a meterstick is 1,000 mm long, but a tiny base-10 unit cube has a volume of $1,000 \text{ mm}^3$ seemed to confuse students. For example, more often than not, students hesitated when asked what superscript to apply to a millimeter measure. The teacher provided a set of problems on the board for students to practice with this kind of unit calculation at the beginning of the last two days of the study, for all three groups. Because of technical difficulties for the COE group, they did their actual online design work after this discussion of volume.

Instrument for Dissertation Study – Estimation Questions

Pencil holder designs were not considered if they violated the volume constraint, and all six of the chosen student designs printed within an hour. The tiny pencil holder described above had a volume of roughly $6,000 \text{ mm}^3$, and printed in just 10 minutes. The teacher designed a very large and elaborate pencil holder that held nine pencils and took three hours to print. This was sent to the 3D printer after the end of the school day.

In addition to engaging in mathematical discussions of volume and the time needed to print various 3D designs, students watched the 3D printer in action. Both of these may have contributed to the phenomenon that was observed during the interviews of students accepting or drastically revising printing time estimates. This is most likely an example of anchoring (Tversky & Kahneman, 1974) based on reference points (Joram et al., 1998), and is examined further in the following chapter.

Results from a Follow-up Interview with the Teacher

The teacher was interviewed two weeks after the unit ended. Several themes emerged from the interview, including: (1) the scope of the project (2) estimation and strategies, (3) uses of the feedback in software, (4) desktop manufacturing possibilities, and (5) possibilities for estimation and 3D software. Selected quotes and summaries of the teacher's reaction follow, organized under these themes.

The teacher offered a mixture of positive reflections on the unit and criticisms of its design. His greatest concern was that it combined computational and volume estimation without fully tying together the skills and strategies involved in each process. These comments fit under the theme of scope.

The Scope of the Project

Regarding the scope of the project, the teacher said the following:

There was a lot of stuff that went on, and a lot of really good stuff that went on, and each individual thing could have received a lot more time. We set a very long list of goals for one week.

He went on to say that the estimation by itself could take multiple weeks, and that at least one or two entire weeks could have been devoted to the use of desktop fabrication. He was concerned by the lack of overlap in the two main themes.

I think there were just two different things going on here, and I think that's one of the things that caused them a little bit of trouble, is, you have the Estimation Calculator, which in its current bit, works with getting accuracy at estimating answers to the basic four operations. And then you have the concept of using the 3D printer and using the 3D design software, where you kind of go to a different field of estimation, where you're asked to come up with some form of units.

Estimation and Strategies

He was concerned by combining computational and volume estimation into one instructional unit. However, he saw value in the separate components. He understood and supported the teaching of estimation, saying "kids are really bad at estimating, which is probably why you're doing research on estimation." He went on to discuss strategies for teaching computational estimation.

When the interviewer mentioned student difficulties with division, he pointed out that "we never really explicitly handled division," and presented a sample division estimation strategy: "You always want to round the larger number in order to have the least effect on it. ...

Instrument for Dissertation Study – Estimation Questions

But maybe that's a little more sophisticated than something they would have tried to do." He referred to manipulation of numerators in fraction estimation as a related example, and one that student generally need practice solving. Computational estimation is generally the domain of elementary school, as the teacher also acknowledged:

When you're dealing with seventh grade, perhaps it's a little old to really spend a significant amount of time dealing with estimating addition of numbers. And I can't really think of somewhere in the curriculum where that would fit. But I definitely think that in fifth or sixth grade, a significant amount of time could go with, 'Hey, look, you should know how to do this by now, but we always want to check to see if we're reasonable.'

Although there is not officially room in the curriculum and pacing for direct teaching or practice with computational estimation in seventh grade, he acknowledged the students in the standard (COE) group did benefit from the practice, saying at one point, "Those standard kids needed a lot of practice with the addition estimation."

He also shared his perspective on methods for teaching volume and related measurement estimation strategies. He brought up the subject of reference points, particularly when doing 3D design on a computer:

I also am very fond of the Google Sketch[up]'s person, that you have standing there to give you an example of the scale of something. And as we were working on pencil holders, it might have been really informative to have a pencil on their screen, so they can see how it looks next to a pencil. As we saw, some kids' pencil holders that were roughly the size of a pen-cap. And some of them we definitely couldn't print because they were the size of a school.

Instrument for Dissertation Study – Estimation Questions

Students' success with volume estimation in the unit was hampered by the by their lack of familiarity with millimeters, which were used extensively in the software. "I think half of them could have told you the number of millimeters in a meter, at the start of the unit. And that's half. Probably three-quarters by the end. They just have very little comfort." He recommended that, in addition to discussing units of millimeters, the English system of measurement would have provided more familiar anchors. Referring to the metric and English systems, he said:

In the world that we live in, they deal with inches, they deal with feet and yards, they deal with pounds, stuff like that. And it is important to go into the metric system, but, somewhere in the discussion, talking about both of them [systems of measurement].

Uses of the Feedback in Software

He acknowledged the COE group's growth in fundamental computational estimation skills, stating

With the standard class, as they were working, and they did improve quite a bit on the addition [estimation] problems. ... They kind of understood the front-end rounding. Some of them understood compensation to a larger degree than others of them. But, for the most part they understood that based on the rounding they did, they – if they were told 'Hey, no, this isn't close enough', they knew kind of to leave the first digit as it was and to look elsewhere.

He noticed a floundering, guess-and-check strategy, that was tied to student frustration:

At first they didn't know how it [the Estimation Calculator] was cool or useful. And then, as they got a little bit more proficient with it, it was like, 'Oh, this is really cool,' and they started being more interested. And then there was also the checkout period, where it was

Instrument for Dissertation Study – Estimation Questions

like, 'Oh, well I've tried fixing it like this and this, and it's still not done, so I don't care anymore.'

This description of student attitudes is informative. It illustrates the teacher's perspective on how students tend to get frustrated and give up rather than pursuing estimation strategies when their estimates are unreasonable. This is likely related to the notably low response rate among students during pre-and post-assessments. Estimation did not seem to be the only area in which students lacked motivation and needed external forces to compel them to work. The teacher was in the process of establishing regular lunch-break-detentions for students who did not turn in their homework. The impact of low student effort was apparent in much of the student work, and is exemplified in student reactions of the kind that the teacher describes.

Desktop Manufacturing Possibilities

Regarding the desktop manufacturing in the unit, the teacher said that "for that [desktop manufacturing] to be more tangible for the kids, it's about finding the right object. I think the 3D printer's tangibility is completely determined by the object that you're printing." He suggested that pencil holders were somewhat abstract for a group of adolescents who did not have a desk job with a regular desk where they kept pencils. His suggestions for possible future desktop manufacturing units included designing cell-phone holders or jewelry, with an emphasis on things that students might use in everyday life.

He felt that the choice of 3D software was useful for motivation ("I think they loved TinkerCAD") albeit with some distractions (such as "all of the letters that most likely were going to have very little impact on the project"). He had interesting suggestions on the use of 3D software to teach geometry:

Instrument for Dissertation Study – Estimation Questions

I do think they enjoyed the 3D printing process, and it was interesting for them to learn a little bit about some of the ways in which it works, and I'm sure they could have done a lot more. Some of them would have been more interested to see – oh, you know, how does it pick what pattern it goes through to do the printing? How exactly does the printing occur? There's a wealth of information that could have been explored with that, that probably would have been more outside the realm of a seventh-grade math class, but is interesting nonetheless.

He went on to make suggestions for future uses of similar software, saying "It would be interesting to have an add-on for TinkerCAD that would allow them to do things and make conjectures about what happened, like, make a cone and then double its height, what's going to happen to its volume?"

Possibilities for Estimation Software

Regarding the use of the WISEngineering estimation step types, he commented that "the estimation stuff that was provided in WISEngineering – I think the fact that it was provided online rather than with pencil and paper had little impact on their motivation to work on it." The feedback provided in the WISEngineering interface was patterned after the Estimation Calculator, with the added interruption after 3 unreasonable estimates. In this way, the error-bar feedback component was interactive and could not have been replicated in pencil-and-paper format. However, the interruption in the interface that required students to reflect upon their strategies could indeed have been provided with paper.

Finally, the teacher discussed creative future solutions for teaching with 3D software, and incorporating estimation:

Instrument for Dissertation Study – Estimation Questions

It would be interesting to see a mathematical interface in the [3D design] software.

Whatever that might look like. Maybe pop in the Estimation Calculator and before they can see what it looks like, they have to estimate its new size.

He also suggested possibilities for interactive online estimation interfaces such as a clickable number-line, and a 3D interface for geometry that might help students derive proofs, such as the formula for the volume of a pyramid, by interacting with solids such as pyramids and prisms.

Summary

The teacher provided insights for the unit that touched on pedagogy and technology. Pedagogically, he suggested, the pace of the unit was too fast. It was designed to cover more material than was feasible for one week's worth of instruction. Features of the students themselves could have been considered as part of the curriculum, such as their lack of experience with metric measurement. In general, the students lacked skill with both computational and measurement estimation. For measurement estimation, the use of virtual reference points, such as a pencil, in the software could have served students in designing more reasonably sized objects.

On the subject of the technology, the teacher suggested more interactivity for the estimation components. He also had some creative uses for desktop manufacturing in the middle school or high school classroom. The desktop fabrication of pencil-holders might have lacked relevancy for the students, but the technology was inherently interesting for them nonetheless. A more careful choice of object, or even a more open-ended design task might have captured student imagination and increased motivation. Some of the features of the 3D printer itself ("How does it pick what pattern it goes through to do the printing?") might require college-level

Instrument for Dissertation Study – Estimation Questions

mathematics to fully grasp, but exploration in a middle school classroom could lay a foundation of understanding for these more advanced concepts.

The follow-up interview was designed to elicit ideas for future study as much as to request a review of the present one. The teacher obliged with a classroom-based perspective and some suggestions for practical applications. These are discussed further at the end of the following chapter.

Outliers

Four student responses appeared as apparent outliers and were eliminated from the analysis because they distorted the distribution of the estimates. The following estimates were outliers. Three of these appear to be spurious, resulting from students misreading the problems, mistyping answers, or simply typing nonsense results. During the final assessment, there was at least one student who verbally admittedly typing nonsense results as estimates.

For the problem: $87419+92765+90045+81974+98102$ (presented as a column of numbers)

The estimate: 25

For the problem: $159 + 164$

The estimate: 20511461

For the problem: $491 - 417$

The estimate: 1232187321

However, the fourth of these drastically unreasonable estimates can be explained using a (flawed) estimation algorithm:

For the problem: $474,257 \div 8,127$

The estimate: 200,000,000

Instrument for Dissertation Study – Estimation Questions

For this estimate, the number of zeroes (8) matches the number of number of non-significant digits that would result from a truncated version of the original problem: $400,000 \div 8,000$. In other words, this is the number of zeroes that a person might retain if he or she were estimating the product, rather than the quotient of the two numbers, with truncation. In that case, the product of $400,000 \times 8,000$ would be 3,200,000,000, which also has 8 zeroes. In this student's estimate, the first digit is 2, which is the quotient of 8 and 4. Thus, though the problem requires dividing 4 by 8, this student appears to have done the reverse. Unfortunately, this student did not provide any explanation of her strategy, so this explanation of the response is speculative. However, it fits with the pattern of student responses in the interview with student COE0402, where strategies for different basic functions were confused. This estimate was made by a student in the COE group who was not either of the interviewees.

Conclusion: Standard Group Student Difficulties

To conclude this chapter, two illustrative examples are given of the causes for unreasonable student estimates. For example, student COE0404 (pre-assessment) used a front end multiplication strategy for the left-most-(tens)-digits and an addition strategy for the ones digits. She provided 125 as an estimate for 62×23 , presumably deriving the first two digits of the estimate (12) by multiplying 6×2 , and then appending the sum of the last two digits, $2 + 3$ as the last digit (5) of her estimate. Student COE0413 (pre-assessment) reformulated $491 + 417$ to $49 + 17$, by simply removing a digit here (final 1 in 491) and a digit there (left-most 4 in 417).

Strategies such as these shed light on the large number of estimates – roughly one third of all student estimates – that fell outside of the reasonable range. These kinds of invented strategies, which tend to confuse place value, are unlikely to result in reasonable estimates. Both

Instrument for Dissertation Study – Estimation Questions

of them violate the preservation of original digits. Direct instruction was given on a multiplication estimation strategy, and this may be responsible for the COE group's incremental movement toward reasonable estimates. Students in the COE group showed particularly strong gains on the multiplication estimation problems, which is exemplified by student COE0402's improved solution of the multiplication post-interview question.

The apparent aptitude-treatment interaction in the simplicity scores for the computational estimation assessment items is likely tied to the fact that many of the COE group students did not apply the principle of simplification (Lefevre et al., 1993) during the pre-interview. In other words, many of these students used exact computation rather than an estimation strategy.

In sum, relative stability of the reasonableness and simplicity scores for the two advanced groups may be attributable to two things. The LAX group received no direct instruction on estimation, while the SAE group did. No movement in scores should be expected from the LAX group. The lack of movement of the SAE students' estimation scores for simplicity and reasonableness may be the result of a ceiling effect. These students' estimation scores were fairly reasonable and mostly simple during the pre-assessment, and did not move. Meanwhile, the COE group's relative (although non-significant) movement on both of these measures appeared as an interaction.

Interviews with students and with the teacher provided extra perspective on these results. The results themselves are summarized in Table 29 at the beginning of the next chapter.

Chapter 5: Discussion

Introduction: Is Estimation Teachable?

The following chapter will address the driving question from the first chapter: *Is estimation teachable?* The specific research questions that fall under this have not been conclusively answered by this study, although the principle of simplification (LeFevre et al., 1993) seems to have had an impact on the estimates of the COE group. Key points from the results are informative for future research, for teachers who wish to teach estimation, and for software developers who intend to develop educational tools.

The present study established three constructs, derived from the literature, to operationalize good estimation. These constructs are reasonableness, simplicity, and nuance. The following table (Table 29) restates the original research questions and addresses each one in turn. Next follow three sections discussing outcomes from the present study, as defined by each of the three main constructs. There is a section on reasonableness, a section on simplicity, and a section on nuance. After that is a brief discussion of the misuse of estimation strategies, which was a recurring theme during the present research and sheds light on future directions for research. The practical implications of the outcomes for teaching, curriculum and software development are treated in the subsequent section. That is followed by a discussion that situates the present research within the existing studies on estimation. The final three sections deal with limitations of the present study, directions for future research, and a conclusion.

Instrument for Dissertation Study – Estimation Questions

Table 29

<i>Summary of Results</i>		
Question	ANOVA	Comment
RQ3) How does reasonableness of estimates differ between students after the two treatments described in R1 [<i>estimation</i>] and R2 [<i>exact computation</i>] for:		<i>Omnibus test, just analyzing reasonableness for COE, SAE, & LAX</i>
a) computational estimation problems?	Time: NSD Group: SD* T × G: NSD	COE made the least reasonable estimates.
b) volume estimation tasks that resemble the problems in the unit?	Time: NSD Group: NSD T × G: NSD	COE estimates were often unreasonable, with slight improvement.
c) multi-step volume estimation problems that do not resemble the problems in the unit?	N / A	Addressed during post-interviews (RQ4).
RQ1) Can instruction on computational and volume estimation combined with automated feedback on estimation in a project-based context lead to students making:		<i>Estimate reasonableness, Strategy complexity, and Strategy nuance analyzed for SAE and COE groups.</i>
d) more reasonable computational and volume estimates?	Mixed Results (see RQ3)	SAE students did not change in reasonableness.
e) simpler computational and volume estimates?	Time: NSD Group: SD* T × G: SD*	COE estimates started complex, with slight simplification.
f) more nuanced computational estimates?	Time: NSD Group: SD* T × G: NSD	SAE and COE students had low median nuance scores.
RQ2) Can an identical project-based context that teaches exact volume computation rather than estimation, and does not include automated feedback lead to students making:		<i>Estimate reasonableness, Strategy complexity, and Strategy nuance analyzed for LAX group.</i>
a) more reasonable computational and volume estimates?	Mixed Results (see RQ3)	LAX students did not change in reasonableness
b) simpler computational and volume estimates?	Time: NSD Group: SD* T × G: SD*	LAX estimates became more complex, did not reach significance
c) more nuanced computational estimates?	Time: NSD Group: SD* T × G: NSD	LAX students made the most nuanced estimates, before and afterward.
RQ4) How do students in the two conditions describe their computational and volume estimation strategies:		<i>Group characteristics emerged:</i>
a) before a project-based unit?	N / A	LAX group: proficient
b) after a project-based unit?	N / A	SAE group: efficient COE group: uncertain
NSD indicates No Significant Difference. The ANOVA results are inconclusive. SD* indicates Significant Differences exist at $p < .05$. These are described in Comments. N / A indicates Not Applicable. ANOVA was not used to address this question.		

Summary of Results

Reasonableness

Research Question 3 can be answered first, since the omnibus analysis of variance (ANOVA) addressed these questions. This research question is phrased so as to address *how* the reasonableness of estimates differs. The result has more to do with group assignment, which was the result of class assignment, than with any changes from pre- to post-assessment.

The construct of reasonableness was defined as how close an estimate is to the exact answer. This is informed by LeFevre et al.'s principle of proximity (1993), which is the idea that an estimate should be close to the exact answer and Reys et al.'s (1982) definition of estimation as providing an approximation that is "reasonably close." It is also informed by Garofalo and Bryant's (1992) discussion of reasonable numbers in the real world, as well as a recent survey of teachers suggesting that the ability to generate reasonable estimates is a core skill and one that remains under recognized in K-12 (Rheinlander et al., 2008). In the present study, reasonableness was operationalized as the distance of an estimate from the exact answer. Reasonableness was measured as absolute magnitude error (AME) mean absolute magnitude error (MAME). This is based on logarithmic transformation of the errors in student estimates, resulting in a measure that roughly captures place value error as a multiple of 1, and front-end digit estimation error as decimal remainder, or distance from the nearest multiple of 1.

The first conclusion that can be drawn is that these seventh students made a large number of unreasonable estimates. Malcolm and Berry's (2011) Estimation Calculator provides a default tolerance for a reasonable estimate as $\pm 15\%$ error, which is a theoretically and pedagogically useful cutoff for reasonableness. After a log transformation, this becomes approximately $\pm .07$. From Figure HISTO, it is clear that about two-thirds of all student estimates cluster close to zero.

Instrument for Dissertation Study – Estimation Questions

The remaining third of the estimates are scattered above and below zero, some as far as four orders of magnitude away. These included computational and volume estimates, and they exclude a handful of outliers that were as far as seven orders of magnitude away from the exact answer.

The results from computational estimation and volume estimation portions of the assessment show that large differences existed between groups both before and after the intervention. The COE group made generally unreasonable estimates before the intervention and slightly more reasonable estimates afterward, which is evident in the downward trend in their estimate errors in Figures 11 and 13. This lessened the gap between their performance and the performance of the other two groups, but not significantly.

There may have been a ceiling effect for the Small Advanced Estimation (SAE) group. (Considering the orientation of Figures 11 and 13, it may be easier to think of this as a floor effect, since their estimates errors did not drop from their initial positions.) Their estimates were relatively reasonable before the unit, and remained relatively reasonable afterward. Students in the SAE group may have been inclined to use more efficient strategies after the intervention than beforehand, as was the case for both students that were interviewed from this group. Any changes in the SAE group's strategies did not greatly influence the resulting estimates.

The Large Advanced Exact (LAX) group was the only group that did not receive direct instruction on any estimation strategies. As Figures 11 and 13 indicate, their estimates became slightly less reasonable, relative to the other two groups, between the beginning and the end of the intervention. This is not a significant change. It could be attributable to a decrease in effort on the post-assessment or even a mere regression to the mean arising from the inherent randomness in their mistakes.

Instrument for Dissertation Study – Estimation Questions

Research Questions 1 and 2 deal with the characteristics of these two groups as measured by the three constructs of reasonableness, simplicity and nuance:

The first part of both questions has been addressed by the analysis for RQ3. Perhaps a similar combination of instructional components could lead to significantly more reasonable estimates for an estimation group, but the implementation in this study did not lead to such a result. While not reaching significance, there was noticeable improvement in the reasonableness of estimates by the Collaborative Estimation (COE) group, which was a standard-level seventh grade mathematics class. There was no perceptible change in the other group receiving the same instruction. This suggests that direct instruction on estimation strategies may be of the more use to standard seventh grade mathematics classes than to advanced mathematics classes. Some of the students in the COE group struggled with the fundamentals of the four basic operations, and could likely have benefitted from more direct instruction on estimation strategies, particularly division estimation strategies. Suggestions for such instruction are included at the end of this chapter.

Before dealing with the subjects of simplicity and nuance in estimation, specific components of volume estimation are discussed here. What follows are subsections on the use of π as a part of volume estimation for cylinders or related shapes, and on the understanding of the context of desktop manufacturing as it informs students' volume estimates.

Use of an approximation for π

Students in both estimation groups improved their estimates for the volume of a cylinder by incorporating an estimate for π into their estimates. While this shift did not reach significance, it lends support to the claim that directly teaching estimation with π improves volume estimation for this kind of problem. The exact computation (LAX) group seemed to

Instrument for Dissertation Study – Estimation Questions

have a good grasp of the use of π before the intervention and their use of it for the cylinder volume estimate was comparable on the pre- and post-assessments. Teaching volume estimation was a goal for the estimation groups, while teaching traditional geometry was the goal for the exact calculation group.

This brings up a definitional question. How does the teaching of volume estimation differ from the teaching of traditional geometry? Clearly, they are interrelated. The difference hinges on Reys et al.'s (1982) definition of estimation as being "reasonably close." Students were shown the formulas to use during both pre- and post-assessments and were also given the suggestion that they use 3 as an estimate for π . Arguably this was an exercise in using and applying geometry formulas, with an emphasis on rounding each operand to simplify the work. The COE and SAE groups made gains in applying and using geometrical formulas that involved the use of π .

Any gain in a core mathematical skill such as this one contributes to gains in the reasonableness of estimates students make within that context. The large overlap between computational and volume estimation in the logic model (Figure 1) represents computational skills that can contribute to reasonable volume estimates when estimating volume. Such skills, which included the ability to read and apply a formula involving π , and to follow the instruction to substitute 3 for π , are necessary but not sufficient for making reasonable volume estimation.

Students will encounter many of the same pitfalls when they estimate volume that they do when they calculate exact answers for volume. Prior research has demonstrated that the difficulty students encounter on any such subcomponent of a geometry problem is more than compounded when such subcomponents are combined (Kao et al., 2007). Volume estimates that the estimation groups, SAE and COE, made during the post-assessment that incorporated π (or

Instrument for Dissertation Study – Estimation Questions

rather, 3) were better than the estimates they made during the pre-assessment. Judging by the place-value errors that students made on problems requiring the use of decimals during the interviews, the choice of 3 as an estimate for π benefits students. Students did not have to concern themselves with the decimal portion of π , which would likely have led to even more confusion and more unreasonable estimates.

In conclusion, the data on the cylinder problem lend support to the notion that an estimation curriculum can improve student estimation performance on specific types of problems. The direct instruction during the unit on estimating with π , as well as practice with this type of problem, may have contributed to this finding. Because the students in the exact calculation (LAX) group showed such strong performance estimating with π at the outset, no comparisons can be drawn about outcomes from the two instructional approaches.

Knowledge of Desktop Manufacturing Constraints

With the exception of student LAX0725, all interviewees accepted their estimates on the pencil-holder problem if they seemed reasonable, and rejected them if they did not. LAX0725 was not prompted to revise his estimate based on a reference point / anchoring strategy, so no conclusion can be drawn about his final estimate. Nevertheless, the other students' responses all fit with the hypothesis that students will make reasonable estimates within a context. Their estimates, calculated and not-calculated, ranged from "50-something" to 99.9 minutes and so are all within an order of magnitude of the actual printing time of the actual pencil holder, which was 45 minutes. In fact, the closest estimate ("50-something minutes") was made by the one student, SAE0317, who did no computation whatsoever.

For this post-interview, students had in their minds a collection of anchors for estimation (Tversky & Kahneman, 1974). More specifically, they had what Joram et al. (1998) refer to as

Instrument for Dissertation Study – Estimation Questions

reference points for measurement estimation. These included three sample pencil holders that were provided at the beginning of the unit, as well as the six small student pencil holders and the large pencil holder designed by the teacher. Discussions in class centered on the volume and print times of each of these pencil holders. Moreover, students witnessed the 3D printer printing the objects, and so established a "feel for" the amount of time it might take to print something out in plastic.

All of this supports Joram et. al.'s (1998) notion that students learn to make measurement estimates based primarily on reference points. In fact, it appears that students may make very reasonable estimates using reference points, even after having gone through a series of steps that leads them to a very different, and less reasonable, conclusion. During the post-interviews, students were more willing to trust their numerical estimates as actual estimates for printing time if the numbers fit within the range to which they had become accustomed. Student LAX0725 failed to identify the proper units for his estimate as minutes, but his result was in the tens of thousands. The girl from the same group did recognize that her estimate of 99.9 should be a measure of minutes. There is interplay between reasonable numbers and reasonable units that seems to be at work in the measurement estimation process.

Simplicity

LeFevre et al.'s (1993) principle of proximity often competes directly with the principle of simplification. Simpler estimates generally involve fewer digits, but may also stray further from the exact answer than estimates involving more digits. For students who seem to lack a "tolerance for error," (Reys et al., 1982) teaching simplicity is probably the higher priority. The first and most important use of computational estimation in everyday life is to provide a check on

Instrument for Dissertation Study – Estimation Questions

values in the real world (Usiskin, 1986). For this, estimation must be easier than exact computation, and greater simplicity makes estimation easier.

Students in the present study received direct instruction on multiplication estimation, using a strategy that emphasizes simplicity. This was a front-end strategy for use with two operands that consisted of the following steps, in this order: (1) rounding, (2) setting aside trailing zeroes, (3), manipulating front-end digits, and (4) appending all trailing zeroes from both operands. Students were encouraged to retain only the left-most digit in each operand, for maximum simplicity. As an analysis technique, non-zero digits that students retained were counted to establish a complexity score. For some problems, complexity of estimates decreased, which is equivalent to greater simplicity.

Simplicity May be Teachable for Computational Estimation

Measuring by this technique revealed a significant group \times time interaction for the computational estimation questions. The COE groups' computational estimates became simpler from pre- to post-assessment (although, by itself, this did not reach statistical significance). Further examination of these values reveals that students in both estimation groups combined made simpler estimates for problems using all four basic operations (+, −, \times , \div), while estimates for the exact (LAX) group became more complex, on average, for all operations except subtraction.

The significant result suggests an aptitude-treatment interaction. Instruction on the simplification principle of estimation had more of an impact on the standard (COE) group's strategies than it did on the advanced (SAE) group's strategies because the principle was largely new to them. Simply put, a sizeable number of the COE students did not know that estimation involves simplifying a problem. On the pre-assessment, more than a third of respondents to the

Instrument for Dissertation Study – Estimation Questions

computational estimation problems simply restated the problem as their reformulation. On the post-assessment, this proportion dropped below 10 percent among COE students.

The movement of the complexity scores for the estimation groups seems to indicate adoption of the rounding algorithm described above and explained by the worksheet. Student SAE0306's increased use of front-end estimation during his post-interview also supports this. For computational estimation, teaching certain specific algorithms such as this one seems to have a direct impact on students' strategy choice and execution.

Greater Complexity May be Desirable for Volume Estimation

On average, all groups showed a slight increase in the complexity of their estimates on the volume estimation problems from pre- to post-assessment. This is sometimes desirable, as the analysis of the use of π indicates. The cylinder estimation problem asks for the volume of a cylinder with a radius of 10 mm and a height of 20 mm. A reasonable reformulation of the formula, $V = \pi \times R^2 \times H$, is $3 \times 10 \times 10 \times 20$ (complexity = 4), where 10×10 is the radius squared. Students in the estimation groups provided reformulations such as 10×20 (an SAE student), $10 \times 10 \times 20$ (from one SAE student and one COE student) or explanations such as "R*H" that reveal a lack of understanding as to how to solve this kind of problem, on the pre-assessment. These reformulations had an average complexity of 3.0. This value went up to 3.25 on the post-assessment, indicating a sufficient number of reformulations with 4 significant digits to bring up the average. This can be directly ascribed to the increased use of 3 as an estimate for π described above. These examples of slightly increased complexity resulting in more reasonable estimates demonstrate a new principle defined for this study, called nuance.

Nuance

Nuance scores are intended to capture what Reys et al. (1982) refer to as processes used by good estimators. Estimate strategies obtained high scores for reaching a kind of middle ground. Strategies that avoided calculating with every last digit (exact calculation), while also avoiding algorithmic rounding (Round-SND keeping only left-most digits), received the highest nuance scores. Reformulations indicate strategies, and reformulations based on techniques such as compensation or compatible-numbers strategies boost a nuance score. The results of this study indicate that students did not learn how to make more nuanced estimates.

There was virtually no movement in nuance scores from pre-assessment to post-assessment for any of the three groups, but there were significant differences between groups. The LAX group received the highest nuance scores for their strategies both before and after the intervention, although these dropped slightly during the post-assessment. This fits with the pattern that the LAX students generally had a good grasp of how to estimate at the outset of the intervention, and though they did not receive any instruction on how to estimate, their strategies remained fairly nuanced afterward. The drop in nuance may indicate a test-retest effect, with two students calculating exact answers for problems they remembered from the pre-assessment and many other students falling back on algorithmic rounding for problems that they were anxious to complete. The LAX interviewees relied on hand-calculation of slightly simplified reformulations during the post-interview, suggesting that proximity influenced their strategies more than simplification. This resulted in relatively high nuance scores.

By contrast, the SAE group generated estimates that were hardly nuanced at all. This fits with their general tendency to be efficient. Both of the interviewees from the SAE group quickly turned to algorithmic rounding strategies, without any use of compensation, and developed

Instrument for Dissertation Study – Estimation Questions

reasonable estimates with minimal writing. Their nuance scores reinforce the conclusion from the simplicity analysis that these students understood and relied on the simplification principle in estimation.

Finally, the COE group received nuance scores that were between the other two groups. Their nuance scores were higher than those of the SAE students but lower than the LAX students. In light of their often unreasonable estimates, it is worth noting that the nuance scores were generated automatically by computer. Thus, illogical reformulations, such as $26 \div 7$ for the problem $426 \div 7$ (nuance = 5), received high scores and raised the median score for the COE group. Nuance should be analyzed after, and in light of, reasonableness and simplicity, in order to keep perspective on the information conveyed by the measure.

The nuance construct was partially developed to investigate students' use of compensation, which is one of the more advanced strategies. Use of compensation was rare among adolescents in the seminal studies (Reys et al, 1980; Reys et al., 1982) and in the present study, it was never seen. As an alternative data source about students' use of compensation, the researchers developed the adjustment checkbox and accompanying text-box for the online assessment. This adjustment feature was used exclusively for place-value adjustment, and was not used at all by either of the estimation groups during the post-assessment. A lesson in the middle of the unit on the use of compensation strategies, which incorporated the Estimation Calculator, did not appear to have great impact on students' thinking about how to make estimates. Data from the adjustment checkbox support this conjecture.

Summary

Between-group differences dominate the results. The COE group's earliest computational estimates suffered from widespread confusion about use of the basic operations, as well as a

Instrument for Dissertation Study – Estimation Questions

tendency to use exact computation in place of estimation. Thus they generally made more unreasonable and more complex computational estimates than those of the other groups. Their estimates became slightly (although not significantly) more reasonable and simpler after the intervention. The apparent aptitude-treatment interaction that resulted from the simplicity analysis is most easily explained by the fact that the COE students, at first, were "keeping all of the digits," while the other groups, by-and-large, were not. At the end of the intervention, the COE students demonstrated a greater tendency to eliminate digits and use what Reys et al. (1982) describe as front-end strategies. Thus, direct instruction on computational estimation strategies may be of the more use to standard seventh grade mathematics classes than to advanced mathematics classes.

Improvements for both estimation treatment groups (SAE and COE) on volume estimation resulted from increased incorporation of π into their reformulations. While neither of these shifts reached significance, it is observable in the reformulations themselves. The other key outcome of the practice with volume estimation in the context of desktop manufacturing is the generally reasonable estimates that students made for the print-time of a pencil holder during post-interviews. This suggests that teaching based on reference-points (Joram et al., 1998) may be fruitful for volume estimation, which could in turn be used as a check on reasonableness of computed estimates.

Misuse of Strategies

Some of the sources of error that caused student estimates to be unreasonable merit further attention. Difficulties with place value were rife among the COE group's estimates, but were common among estimates of the other two groups as well. This was sometimes caused by

a confusion of strategies. Particularly among the COE students, techniques that are valid for one of the four basic operations were often applied to the other three operations, such as multiplication techniques applied to division problems. For example, the student who provided the estimate 200,000,000 for the problem $474,257 \div 8,127$ seems to have used the "keep all of the zeroes" technique that was taught for multiplication. This response was given by one of the COE students during the post-assessment. She was apparently misapplying what she had learned about multiplication estimation. For a future study or teaching with a group demonstrating this much difficulty with basic operations, direct instruction as well as practice on division estimation could serve to improve students' division estimation reasonableness.

Implications: Relevance of the Findings

The present research has implications for teachers, for writers of curriculum, and for software developers, as well as for researchers. The implications for future research are at the end of this chapter. The following three subsections describe the salient points that have emerged for the former three categories of professionals: teachers, as well as curriculum and software developers.

Teaching

Three important points for teaching can be made based on the outcomes of this study. These are specifically applicable to middle-school mathematics teaching, and they center on teaching fundamentals of mathematics, even when students are at a grade level where they "should" know these fundamentals.

Sometimes middle school students lack the basic strategies. The teacher's comments during the follow-up interview encapsulate the dilemma of teaching computational estimation in

Instrument for Dissertation Study – Estimation Questions

middle school. On the one hand, as he points out, "kids are really bad at estimating". On the other hand, "when you're dealing with seventh grade, perhaps it's a little old to really spend a significant amount of time dealing with estimating addition of numbers". Addition estimation, as well as estimation with the other three basic functions, is covered in upper elementary school. However, for students who have never established the number sense and related abilities (Rubenstein, 1985) to become proficient estimators, this kind of foundational mathematics study is crucial (Rheinlander et al., 2008).

Teaching the basic strategies influences outcomes. One conclusion that can be drawn from this research is that, for students who lack these fundamental skills, direct instruction can make a difference. Some of the outlying, unreasonable estimates on the post-assessments were attributable to students making estimates based on the multiplication estimation strategies they had just learned. Meanwhile, students' multiplication estimates tended to improve, which can be directly attributed to the fact that students were taught multiplication estimation strategies. Therefore, while it may seem self-evident, this research has shown that students will use the techniques that they have just been taught. The less obvious conclusion is that sometimes middle-school students need to be taught, or re-taught, fundamental techniques for solving problems with the basic four operations.

Different groups have different needs. The measurable results from this study all originate in the differences between groups. The COE group's estimates became slightly more reasonable, and a large number of COE students shifted from an exact computation strategy to a simpler strategy from pre-assessment to post-assessment. The SAE group did not seem to benefit greatly from estimation instruction and practice opportunities, and neither their computational estimates nor the exact-calculation (LAX) group's computational estimates shifted

Instrument for Dissertation Study – Estimation Questions

in terms of reasonableness, simplicity or nuance. Meanwhile, the COE group could still have benefitted from further practice distinguishing between the necessary strategies on the basic functions.

Curriculum Development

For developers of middle school mathematics curriculum, some of the same key points apply. For teachers, a sensitivity to students' lack of knowledge of fundamental uses of computation strategies are valuable, as discussed above. The following suggestions are provided for specific areas where emphasis could be placed in planning and creating curriculum.

Middle school students can benefit from instruction on place value and computational strategies. One simple suggestion for developers of curriculum is to teach specific estimation strategies. The following division estimation axiom was not taught, but could have saved time for four of the six interviewees during the post-interviews:

Eliminating an equal number of zeroes from the ends of both the divisor and the dividend and then manipulating the front-end digits greatly simplifies division estimation.

In other words:

$60,000 \div 1,000$ is analogous to $60x \div 1x$, where x is any value.

In other words:

An estimate for $60,000 \div 1,000$ is the same as an estimate for $60 \div 1$.

A robust division estimation algorithm, then, can start with rounding (using Rounding-SND), followed by simply eliminating an equal number of zeroes from the right-hand side of both numbers. Strategies such as this one may be simple to teach, and will be useful for students' development of reasonable estimates.

Instrument for Dissertation Study – Estimation Questions

The difficult part of the curriculum is drawing the connections. Students gained an intuitive understanding of how long a 3D printer takes to print an object, and could make reasonable estimates for this almost instantly when asked about a given 3D object. However, drawing the connection between this estimate and the computations it might take to figure this out on paper was much trickier. Complexities of the process got in students' way and the lack of fundamentals described above became a stumbling block. Ideally, the intuitive estimates could provide an anchor for students to work backwards, so that they could reassess computed estimates in light of reality, and vice versa.

Sophisticated estimation strategies such as compensation may not be worth teaching as part of the standard middle-school mathematics curriculum. These strategies might be best saved as enrichment exercises. They seemed to overwhelm students who were already devoting considerable cognitive resources to the processes of rounding and managing digits. The present research found that no students adopted compensation as a strategy. Students understood that some multiplication estimates will be too low as a direct result of having rounded operands down, for example. However, adding a step to the estimation process to compensate for this did not seem manageable. For students who already have plenty of standards-based content to learn, compensation in estimation may be too much to expect.

Keep curriculum simple. Be careful of overwhelming students (and teachers) with too many ambitious goals. As the teacher pointed out, this study involved teaching computational and volume estimation, both within a matter of a week. In the teacher's own words, "each individual thing could have received a lot more time." Results from this study suggest that at least twice as much time could have been spent on either subject.

Instrument for Dissertation Study – Estimation Questions

Project-based instruction takes time. This result echoes Blumenfeld et al.'s (1991) framework for PBL. The present study borrowed philosophical features from PBL, but by no means did it seek to instantiate all of project-based instruction. As a project for grounding mathematics, the design and creation of pencil-holders or other everyday objects holds promise within a broader digital fabrication initiative (Berry et al., 2010). In the present study, the novelty of the 3D printer and the lush interactivity of TinkerCAD seemed to motivate students, and some students drew the connections between the design activity and the underlying mathematics. However, much more time could have been spent ensuring that all of the students made these connections. Given enough time, students could have been tasked with generating driving mathematical questions themselves.

The mechanics of the desktop manufacturing hold further possibilities for learning. Students expressed interest in the functioning of the 3D printer, which could have been the basis for a rich discussion on the meaning of the three dimensions, which in turn could have motivated an understanding of the superscripts on units such as mm^3 and in^3 . These ideas were inspired by bringing a printer into a classroom, watching students' reactions to it, and reconsidering the learning goals in the curriculum. Curriculum will improve as curriculum developers exploit these natural connections. The teacher elaborated on still more sophisticated questions involving the 3D printer: "How does it pick what pattern it goes through to do the printing? How exactly does the printing occur?" These questions concern both mathematics and physics, and connect to curricular content for high school and beyond.

Software Development

One of the components of interest in this curriculum was the students' use of new software for learning, particularly the software components that provided feedback. The following paragraphs list implications for future software development involving estimation.

Rich sources of data for researchers differ from sources of motivation for students. The low student response rate during assessment phases confirms the teacher's comments about student frustration and low motivation. Because of the number of unreasonable attempts made by students during the practice exercises, the practice exercises resulted in students spending most of their time explaining themselves on the interruption screen interfaces. To an observer it was difficult to see the interactive component of WISEngineering at work, since students made so many unreasonable estimates that, by and large, the interruption screen prompting students to reflect on their work was all that could be seen during these sessions. This prompt screen was the same as the assessment screen that was the core of the assessment, both pre- and post-intervention, and was interchangeable with a paper assignment. The prompts on the screen were intended to promote metacognition, and metacognition may have had an influence on the slight shift in students' strategies. Students' responses to the prompts were also central to the analysis in the present study. In short, the interrupt-and-prompt-for-explanation design pattern of software seems to be useful as the basis for research, but deflates motivation for students.

The analyses used in the present study have potential as formative assessment tools. Measuring estimate reasonableness based on the logarithmic transformation illustrated in the previous chapter has proven to be a useful technique for the present research. It allows for the quick identification unreasonable estimates based on the number of "missing zeroes" or "extra zeroes" in a student's estimate. This enables a researcher to compare order-of-magnitude

overestimates and underestimates on roughly the same scale. It could also serve as the basis for automated analysis of student strategies, including generating real-time, intelligent feedback for students on their estimates. Detection of order-of-magnitude errors with a log transformation is more precise than other algorithms such as digit counting, although direct analysis of the digits students have entered shows promise too, particularly if students are asked to enter reformulations. The amount of data requested from students in this study was perhaps cumbersome. While it was useful to have written descriptions of strategies when reformulations were garbled or unclear, the reformulations provided in numbers and symbols were generally sufficient. For future work with dynamic feedback on estimation, asking students to briefly type their reformulations seems worthwhile.

Numerous possibilities exist for interactive estimation and 3D mathematics software.

During the follow-up interview, the teacher suggested a handful of interesting ideas for interactive mathematics software. Among these suggestions were 3D tools for deriving geometric proofs, and educational CAD software that require students to calculate or estimate the result that scaling a figure will have on its volume. As a guiding activity for a unit on desktop fabrication, he proposes for student-created 3D artifacts.

Extending Prior Research

This study builds on prior research in the areas of computational and measurement estimation. In addition to prior work with the Estimation Calculator (Malcolm & Berry, 2011), this study addresses specific calls to action in key precursors in the related literature. The first of these is Reys et al.'s (1980) report, entitled *Identification and Characterization of Computational Estimation Processes Used by In-School Pupils and Out-of-School Adults*. The end of this

Instrument for Dissertation Study – Estimation Questions

seminal report includes a list of 6 questions, two of which are paraphrased here. One question (#4) is a call to perform more research on "sensitivity to unreasonable answers and techniques used to identify out-of-the-ballpark answers."

Techniques for teaching estimation and giving students practice with estimate feedback were the central concerns of the present research. Students' sensitivity to unreasonable answers seems to depend on their mathematics background. The groups of students in the advanced classes tended to be sensitive to the reasonableness of answers to a greater extent than students in the standard class. As a technique to identify "out-of-the-ballpark" answers, the use of a physical pencil holder during the follow-up interview provided a check on the unreasonable estimates for time needed to 3D print an object. Students tended to revise unreasonable estimates very quickly when the question regarding printing time was framed with the object, rather than as the result of a series of computational estimates.

Reys et al.'s (1980) sixth and final question is actually a recommendation to develop: ... a systematic plan for teaching estimation to students in grades 7-12 including the most prevalent and effective techniques and strategies identified by this investigation. ... Can students with only minimum estimation ability be taught to use strategies identified among good estimators? (p. 257)

The present study is presented as a first step toward building curriculum for systematically teaching estimation using some of the strategies identified in the Reys report. On a trajectory of learning, it is more advanced than a fourth grade estimation curriculum (e.g., Lan et al., 2010) but does not need to be as difficult as a curriculum for pre-service teachers (e.g., Bestgen et al., 1980; Fung and Latulippe, 2010).

Instrument for Dissertation Study – Estimation Questions

In Reys et al.'s (1980) question regarding whether students with minimum estimation ability can be taught to use the more advanced strategies of good estimators, the data indicate that for such students, one week is not sufficient. The strategy of compensation is a centerpiece of good estimation, but may be too advanced to merit the time it would take away from other subject to teach. Other, simpler strategies such as techniques for division estimation may prove valuable however.

LeFevre et al. (1993) defined the principles of simplification and proximity in the context of multiplication estimation. The same principles apply for other kinds of estimation as well. The present study builds upon this work by developing measures for proximity (i.e., reasonableness) and simplification, and applying them to student estimates. The principle of simplification seems to be teachable. Over one third of respondents from the COE group used exact calculation on the pre-assessment when they were asked to estimate. Less than a tenth of them did on the post assessment. This contributed to a significant group \times time interaction for the complexity scores, with the COE group providing simpler estimates on the post-assessment by comparison with the other two groups.

One method for simplifying the estimation process involves using anchors for estimation (Tversky & Kahneman, 1974). This does not ensure proximity, or reasonableness, of the resulting estimate unless the anchors are chosen appropriately. The anchors, or reference points (Joram et al, 1998), that students were given did not seem to have a large impact on their ability to make volume estimates. Instead, the mechanics of calculating the volumes of cylinders using π seemed valuable to students and did have somewhat of an influence on the reasonableness of their volume estimates. For the physical objects, including pencil holders and base-10 blocks that students were given as anchors, their lack of familiarity with metric units may have hindered

their use of these as reference points. However, students exhibited an intuitive ability to translate between perceived volume and a reasonable estimate for 3D printing time.

Joram et al. (1998) describe a two way process "from number to quantity and back." This two-way translation of measurement estimation was a stumbling block for students, and future studies on volume estimation can focus more directly on translating estimates in both directions. An understanding of the relationship between a unit of length, such as millimeters or inches, and unit of volume, such as mm^3 or in^3 , is necessary for forming good volume estimates. Students struggled with the units, as the teacher pointed out during the interview, including the superscripts that accompanied the units. The fact that a meter stick is 1,000 mm long, but a tiny base-10 unit cube has a volume of $1,000 \text{ mm}^3$ seemed to confuse students. Further time could have been spent discussing the concepts of space-filling with cubes, and how that relates to the mathematical concept of cubing a number.

Berry et al. (2010) proposed the use of digital fabrication in the elementary mathematics curriculum. The present research provides some indicators of how best to approach such research. There are natural connections between DF, as embodied in desktop manufacturing, and the teaching of basic geometry. A fundamental understanding of geometric solids and their volume, as it features in the seventh grade mathematics curriculum, is first provided to students in the upper elementary grades. These concepts featured in the present research as the foundations for creating volume estimates for 3D designs.

Limitations

The quasi-experimental nature and small sample size of the present study limit its generalizability. True randomized assignment into treatment conditions would mitigate the

Instrument for Dissertation Study – Estimation Questions

striking differences between groups that were apparent here. Differences in ability between students in standard and students in advanced-level middle school mathematics class became a key variable in this study, by virtue of pre-existing classroom assignment, rather than by design.

Given this split nature of mathematics class assignment by ability, a better approximation of a truly crossed design would be possible. That is, this study included a standard-estimation group (COE), an advanced-estimation group (SAE), and an advanced-exact group (LAX). A fourth, standard-exact group would have provided a useful contrast to the COE group and would have completed a crossed-design. A future study like this might help establish whether any shift in COE estimation scores was simply due to practice effects of the assessment, for example.

All data was collected, organized and analyzed by the lead researcher on the study. Care was taken to review details and avoid biases in each step of the analysis process, but flaws or subconscious biases may have arisen. The data that was collected was also limited by a response bias, so that the data on any item on the pre- or post-assessment tell the story of only the students who chose to respond.

Directions for Future Research

The present research suggests several directions for future study. These fall under the categories of estimation and digital fabrication. While a future study might seek to combine these two themes, the present study has indicated that the time required to fully integrate the central themes of both topics might be prohibitive. Nevertheless, themes from the present study are ripe for deeper investigation.

Estimation Research

First, the students' usage patterns in interactive estimation activities with WISEngineering were collected as part of this research. Analysis was outside the scope of the present study. A closer look at this data would address the question, Does the disruptive feedback alter the manner in which students respond to iterative estimation problems? Analysis would reveal if estimates students made were based on feedback, and how this was influenced by the prompts to describe their strategies with words and symbols. It would also inform whether students made more reasonable estimates on successive uses of the interactive step-type.

The Estimation Calculator and a WISEngineering step-type with interactive error-based feedback were used as components of the intervention in this study. The small sample size and brief nature of the intervention limit its generalizability. A longitudinal study with a larger sample size based solely on the use of one error-based estimation interface such as the Estimation Calculator would provide valuable data about possible improvements in student outcomes after extended use of this kind of tool. For a study of this nature, an upper-elementary student population would be more appropriate than students in middle school. A possible extension is to include more informative, real-time feedback to diagnose student errors and provide hints after repeated unreasonable estimates. The results of the present study suggest that direct instruction of estimation strategies would be necessary before or during students' use an estimation tool to ensure improvement.

Further research is warranted on the impact of direct instruction of specific estimation strategies that were absent here, such as the elimination of zeroes to simplify division, or the concept of only keeping the large operand during addition or subtraction estimation when operands have different magnitudes. A handful of useful rules-of-thumb for estimation (e.g.,

Instrument for Dissertation Study – Estimation Questions

Rose & Morgan, 1986; Wheatley & Hersberger, 1986) might provide the basis for more reasonable, simpler, and more nuanced estimates among upper elementary students. This would strengthen the foundation of estimation skills, so that students would arrive in middle school more confident in their estimation abilities.

Confidence itself may be an important component of estimation, and is undoubtedly entwined with the concept of tolerance for error. Further study into building confidence in estimates would provide answers as to how to help children become more proficient and more efficient estimators. Multiple instances of inefficiency that were related to a lack of confidence in estimation ability arose during this study. Helping students achieve estimates more quickly by way of practice, and the building of self-efficacy (Bandura, 1977) is worthy of further study.

Research into Digital Fabrication and Desktop Manufacturing

The results of this highlight the rich possibilities for studies that focus on an understanding of proportionality and scale in 3D software among seventh grade students. Software could be developed that provides dynamic feedback on volume problems based on students' shapes and designs. As the teacher suggested in the follow-up interview, this software might restrict a user's actions in software until that user had provided, for example, a reasonable estimate for the stretched or scaled volume of a design on the screen. The impact of this on student's understanding of scale, ratio and proportion would provide useful insights for educators and software developers.

Research that focuses on 3D design and fabrication as a vehicle for geometry holds promise as well. Research in this area has suggested that significant middle school mathematics learning gains are possible with the proper use of engineering design pedagogy and curriculum development (DeJaegher et al., 2012).

Instrument for Dissertation Study – Estimation Questions

Summary

In sum, the possibilities for future research can investigate the impact of estimation processes and tools, as well as desktop manufacturing processes and tools. Interactive software might help integrate these two. As suggested in earlier chapters, the direct instruction of mathematics fundamentals and processes is a pre-requisite for practice and exploration with such tools. The influence of teaching simple techniques to improve estimates requires further investigation, as does the impact of interweaving mathematics pedagogy with digital fabrication and desktop manufacturing.

Conclusion

The application of the simplification principle (LeFevre et al., 1993) for estimation is the central theme of these outcomes. Ironically, application of a similar principle of simplification in the development of curriculum might have benefitted all of the participants in the study as well. The changes that occurred between pre- and post-assessment estimates for the standard class of students seemed to result directly from the most direct and simple parts of the curriculum. When students were given simple directions to follow to generate better multiplication estimates, the students followed these directions. The term *better* is chosen carefully here; a good estimate is one that is both simple and reasonable, and the standard class' final estimates seemed to exhibit both qualities in ways that their initial estimates did not, although such improvements did not reach significance.

The most significant result was the apparent difference in impact on student strategies resulting from group membership (i.e., class assignment). An aptitude-treatment-interaction for estimation instruction in middle school indicates that advanced students do not need help with

Instrument for Dissertation Study – Estimation Questions

estimation, while standard class students might benefit from instruction and practice estimating. This interaction in simplicity (complexity) scores highlighted the fact that students in the standard class used exact computation heavily during the pre-assessment. By the end of the intervention, these students were less likely to rely on exact computation, suggesting a greater tolerance for error and a broader understanding of estimation itself. The vast majority of advanced students knew to reformulate the original problem to form an estimate, both before and after the intervention.

The desktop manufacturing component of the intervention showed promise as well, although the curriculum requires revision. More time would need to be spent laying groundwork on concepts such as units to have a significant impact on student outcomes for volume estimation or exact volume calculation. The possibilities of using 3D software to teach 3D geometry are many, and this study explored just a few of these. As free and inexpensive software tools such as simple CAD programs become more widely available, and more educational, this area will undoubtedly blossom.

Meanwhile, the possibilities for estimation practice software are clearer in light of this study. Some of the post-hoc analysis techniques on students' estimates and strategies provide valuable information that could be used to provide more informative feedback. A feedback bar for the reasonableness of an estimate has shown educational promise. This would be more powerful still when combined with automatically generated hints on strategy improvement.

More than 30 years elapsed between Reys et al.'s (1980) seminal study in computational estimation, which resulted in the conclusion that children and adolescents need help improving their estimation strategies. As the series of studies described in this document demonstrate, there is still work to be done.

Instrument for Dissertation Study – Estimation Questions

References

- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61, 315-343.
- Allen, G. L., & Rashotte, M. A. (2006). Training metric accuracy in distance estimation skill: pictures versus words. *Applied Cognitive Psychology*, 20(2), 171–186.
doi:10.1002/acp.1174
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole Pub. Co.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Anderson, J. R., Boyle, C. F., Corbett, A., & Lewis, M. W. (1990) Cognitive modelling and intelligent tutoring. *Artificial Intelligence*, 42, (pp. 7-49).
- Anderson, J. R., Boyle, C. F., & Yost, G. (1986) The geometry tutor. *The Journal of Mathematical Behavior*, 5-20.
- Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP Tutor. *Cognitive Science*, 13, 467-506.
- Anderson, J. R., & Pelletier, R. (1991). A development system for model– tracing tutors. In *Proceedings of the International Conference of the Learning Sciences* (pp. 1–8). Evanston, IL. Retrieved from <http://act-r.psy.cmu.edu/papers/33/JRA-RP-ModelT.pdf>
- Anderson, J.R., Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167-207
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279–1333.
- Baddeley, A. D. (1986). *Working memory*. Oxford Psychology Series, No. 11. Oxfordshire: Clarendon Press.
- Baker, R.S.J.d., de Carvalho, A.M.J.A., Raspat, J., Aleven, V., Corbett, A.T., Koedinger, K.R. (2009) Educational software features that encourage and discourage "gaming the system". *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, 475-482.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/847061>
- Barnes, T. (2005). The q-matrix method: Mining student response data for knowledge. *Proceedings of the AAAI-2005 Workshop on Educational Data Mining*, July 9-13, 2005, Pittsburgh, PA.
- Bell, B. L., Brown, A., Bull, G., Conly, K., Johnson, L., Mcanear, ... Sprague, D., (2010). A special editorial: Digital fabrication revolution, *TechTrends* 54(5).
- Berry, R. Q., III (1998). Computational estimation skills of eight grade students. (Unpublished master's thesis). Christopher Newport University, Newport News, VA.
- Berry, R. Q., III, Bull, G., Browning, C., Thomas, C. D., Starkweather, K., & Aylor, J. H. (2010). Preliminary considerations regarding use of digital fabrication to incorporate

- engineering design principles in elementary mathematics education. *Contemporary Issues in Technology and Teacher Education*, 10(2). Retrieved from <http://www.citejournal.org/vol10/iss2/editorial/article1.cfm>
- Bestgen, B. J., Reys, R. E., Rybolt, J. F., & Wyatt, J. W. (1980). Effectiveness of systematic instruction on attitudes and computational estimation skills of preservice elementary teachers. *Journal for Research in Mathematics Education*, 11(2), 124–136. Retrieved from <http://www.jstor.org/stable/748904>
- Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2010). Creating Usable Innovations in Systemic Reform : Scaling Up Technology-Embedded Project-Based Science in Urban Schools. *Educational Psychologist*, 35(3), 149–164.
- Blumenfeld, P., Soloway, E., Marx, R., Krajcik, J., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26, 369-398.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: The National Academies Press. Retrieved from <http://www.nap.edu/>
- Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: A framework for understanding real-world quantitative estimation. *Psychological Review*, 100(3), 511–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8356188>
- Bull, G., & Groves, J. (2009). The democratization of production. *Learning and Leading with Technology*, 37(3), 36-37.
- Burghardt, M. D., & Hacker, M. (2004). Informed design: A contemporary approach to design pedagogy as the core process in technology. *Technology Teacher*, 64(1). 6.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. doi:10.3102/00346543065003245
- Case, R., & Sowder, J. T. (1990). The development of computational estimation: A neo-Piagetian analysis. *Cognition and Instruction*, 7(2), 79-104.
- Chi, M. T. H. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Erlbaum. Reprinted in: Wozniak, R. H. (1993) *Worlds of Childhood*, (pp. 232-240), Harper Collins College Publishers. Retrieved from <http://chilab.asu.edu/papers/Chi.pdf>
- Chiu, J. L. & Linn, M. C. (2011). Knowledge integration and WISE engineering. *Journal of Pre-college Engineering Education Research*, 1(1), 1-14.
- Clariana, R. B., Wagner, D., & Roher Murphy, L. C. (2000). Applying a connectionist description of feedback timing. *Educational Technology Research and Development*, 48(3), 5–21.
- Corbalan, G., Paas, F., & Cuypers, H. (2010). Computer-based feedback in linear algebra: Effects on transfer performance and motivation. *Computers & Education*, 55, 692-703.
- Corbett, A. T., & Anderson, J. R. (1991). Feedback control and learning to program with the CMU LISP tutor. *Proceedings of the 1991 Annual Meeting of the American Educational Research Association*. Chicago, IL: AERA. Retrieved from <http://act-r.psy.cmu.edu/papers/>

Instrument for Dissertation Study – Estimation Questions

- Corbett, A. T. & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of ACM CHI'2001 Conference on Human Factors in Computing Systems*, 245-252.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-114.
- DeJaegher, C., Chiu, J. L., Burghardt, M. D., Hecht, D., Malcolm, P., & Pan, E. (2012). Learning common core math concepts with WISEngineering. *Proceedings paper presented at the American Society for Engineering Education Conference, ASEE 2012*, (June 2012, San Antonio, TX)
- Dowker, A. (1992). Computational estimation strategies of professional mathematicians. *Journal for Research in Mathematics Education*, 23(1), 45-55.
- Dowker, A., Flood, A., Griffiths, H., Harriss, L., & Hook, L. (1996). Estimation Strategies of Four Groups. *Mathematical Cognition*, 2(2), 113–135.
- Dunn-Rankin, D. (2001). Evaluating design alternatives – the role of simple engineering analysis and estimation. *2001 ASEE Annual Conference & Exposition*.
- Earley, P. C., Northcraft, G. B., Lee, C., & Lituchy, T. R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. *Academy of Management Journal*, 33(1), 87–105.
- Easley, D., López de Prado, M. M., & O'Hara, M. (2010). Microstructure of the 'Flash Crash': Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading". *The Journal of Portfolio Management*, 37(2), 118–128.
- Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, 34(5), 433–463. Retrieved from <http://www.jstor.org/stable/30034795>
- Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, 208(4448), 1181–1182. Retrieved from <http://www.jstor.org/stable/1683736>
- Feigenson, L., S. Dehaene, and E. Spelke. (2004). Core systems of number. *TRENDS in Cognitive Sciences*, 8(7). Retrieved from: <http://www.sciencedirect.com>
- Fung, M. G., & Latulippe, C. L. (2010). Computational estimation: Preservice teachers learn how to ground their own number sense. *Teaching Children Mathematics*, (October), 170–176.
- Garofalo, J. & Bryant, J. (1992). Assessing reasonableness: Some observations and suggestions. *Arithmetic Teacher*, 40, 4, 210-12.
- Gershenfeld, N. A. (2005). *Fab: The coming revolution on your desktop--from personal computers to personal fabrication*. New York: Basic Books.
- Gorard, S. (2004). Revisiting a 90-year-old debate: the advantages of the mean deviation. *Paper presented at the British Educational Research Association Annual Conference, University of Manchester, 16-18 September 2004*. Accessed online at: <http://www.leeds.ac.uk/educol/documents/00003759.htm>
- Guay, R. B. (1976). *Purdue spatial visualization test*. Educational Testing Service Test Collection. Princeton, NJ: ETS
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park: CA: Sage

- Halabi, A. (2006). Applying an instructional learning efficiency model to determine the most efficient feedback for teaching introductory accounting. *Global Perspectives on Accounting Education*, 3, 93-113
- Hanson, S. A., & Hogan, T. P. (2000). Computational estimation skill of college students. *Journal for Research in Mathematics Education*, 31(4), 483–499.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, 28(4), 603-621.
- Hogan, T. P., & Brezinski, K. L. (2003). Quantitative estimation: One, two, or three abilities? *Mathematical Thinking and Learning*, 5(4), 259–280. doi:10.1207/S15327833MTL0504_02
- Hogan, T. P., Wyckoff, L. A., Krebs, P., Jones, W., & Fitzgerald, M. P. (2004). Tolerance for error and computational estimation ability. *Psychological Reports*, 94(3), 1393–1403. Retrieved from <http://www.amsciepub.com/doi/abs/10.2466/pr0.94.3c.1393-1403>
- Hope, J. (1986). Mental calculation: anachronism or basic skill?. In H. L. Schoen & M. J. Zweng (Eds.). *Estimation and mental computation: 1986 yearbook* (pp. 45-54). Reston, VA: National Council of Teachers of Mathematics.
- Howie, E., Sy, S., Ford, L., & Vicente, K. J. (2000). Human-computer interface design can reduce misperceptions of feedback. *System Dynamics Review*, 16(3), 151–171.
- Imbo, I., Duverne, S., & Lemaire, P. (2007). Working memory, strategy execution, and strategy selection in mental arithmetic. *Quarterly journal of experimental psychology* (2006), 60(9), 1246–64. doi:10.1080/17470210600943419
- Jonassen, D. H. (2003). *Learning to solve problems with technology: A constructivist perspective*. Upper Saddle River, N.J: Merrill.
- Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research*, 68(4), 413–449.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kalaman, D. A., & LeFevre, J.-A. (2007). Working memory demands of exact and approximate addition. *European Journal of Cognitive Psychology*, 19(2), 187–212. doi:10.1080/09541440600713445
- Katehi, L., Pearson, G., & Feder, M. (2009). *Engineering in K-12 Education*. Washington, DC: The National Academies Press.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist* 41(2) 75-86.
- Koedinger, K. R., McLaughlin, E. A., & Stamper, J. C. (2012). Automated student model improvement. In K. Yacef, O. Zaïane, H. HersHKovitz, M. Yudelson, and J. Stamper, (Eds.) *Proceedings of the 5th International Conference on Educational Data Mining (EDM 2012)* Chania, Greece, June 19-21 2012. 17–24.
- Koedinger, K. R., & Anderson, J. R. (1993). Reifying implicit planning in geometry: Guidelines for model-based intelligent tutoring system design. In S. Lajoie & S. Derry (Eds.), *Computers as cognitive tools*. Hillsdale, NJ: Erlbaum.

- Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8, 30–43.
- Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cognitive science*, 36(5), 757–98. doi:10.1111/j.1551-6709.2012.01245.x
- Koedinger, K. R., & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, 19(3), 239–264. doi:10.1007/s10648-007-9049-0
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254–284.
- Lan, Y.-J., Sung, Y.-T., Tan, N., Lin, C.-P., & Chang, K.-E. (2010). Mobile-device-supported problem-based computational estimation instruction for elementary school students. *The Journal of Educational Technology and Society*, 13(3), 55–69.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child development*, 78(6), 1723–43. doi:10.1111/j.1467-8624.2007.01087.x
- LeBoeuf, R., & Shafir, E. (2006). The long and short of it: physical anchoring effects. *Journal of Behavioral Decision Making*, 19(4), 393–406. doi:10.1002/bdm.535. Retrieved from <http://onlinelibrary.wiley.com>
- LeFevre, J., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computational estimation. *Cognition and Instruction*, 11(2), 95–132.
- Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development*, 26(3), 282–294. doi:10.1016/j.cogdev.2011.01.002
- Levine, D., (1982). Strategy use and estimation ability of college students. *Journal for Research in Mathematics Education*, 13 (5), 350–59.
- Linn, M.C., Lee, H. –S., Tinker, R., Husic, F., & Chiu, J.L. (2006). Teaching and assessing knowledge integration in science. *Science*, 313(5790), 1049-1050.
- Lomax, R.G. (2007). *An Introduction to Statistical Concepts (2nd ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Maier, D. (1978). The complexity of some problems on subsequences and supersequences. *Journal of the ACM (JACM)*, 25(2), 322-336.
- Malcolm, P. & Berry, R. Q., (2011). The estimation calculator. *Teaching Children Mathematics*, 18. 119- 124. Reston, VA: National Council of Teachers of Mathematics (NCTM).
- Malcolm, P., Chiu, J. L., Pan, E., & Burghardt, M. D. (2012). WISEngineering: A web-based engineering design learning environment. *Proceedings paper presented at the American Society for Engineering Education Conference, ASEE 2012*, (June 2012, San Antonio, TX)

- Mathan, S., (2003). *Recasting the feedback debate: Benefits of tutoring error detection and correction skills*. (Doctoral dissertation). Retrieved from CiteSeerX. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.67.8448&rep=rep1&type=pdf>
- Mathan, S. & Koedinger, K. R. (2005). Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational Psychologist*, 40(4), 257–265. doi:10.1207/s15326985ep4004_7
- McKendree, J. E. (1990). Effective feedback content for tutoring complex skills. *Human Computer Interaction*, 5, 381–414.
- Meyer, J. P., & Seaman, M. A. (2013). A comparison of the exact Kruskal-Wallis distribution to asymptotic approximations for all sample sizes up to 105. *The Journal of Experimental Education*, 81(2), 139–156. doi:10.1080/00220973.2012.699904
- Merriam-Webster. (2012). Tolerance (definition). Retrieved from: <http://www.merriam-webster.com/dictionary/tolerance>
- Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Milson, R., Lewis, M. W., & Anderson, J. R. (1990). The teacher's apprentice: Building an algebra tutor (pp.53-71). In R. Freedle (Ed.), *Artificial Intelligence and the Future of Testing*. Hillsdale, NJ: Erlbaum.
- Nathan, M. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments*, 5,(1). 135-159.
- National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics [Online]. Available: <http://www.nctm.org/standards/content.aspx?id=26792>
- National Council of Teachers of Mathematics. (2006). *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics*. Reston, VA: The National Council of Teachers of Mathematics, Inc.
- National Governors Association, (2010a). Common core state standards for mathematics. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C. Retrieved from <http://www.corestandards.org>
- National Governors Association, (2010b). Common core state standards for mathematical practice. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C. Retrieved from <http://www.corestandards.org>
- Neves, D. M., & Anderson, J. R. (1980). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57-83). Hillsdale, NJ: Erlbaum.
- Newell, A., & Rosenbloom, P. S. (1980). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.
- Norris, S. P., Macnab, J. S., & Phillips, L. M. (2007) Cognitive models of performance on diagnostic achievement tests. In J. Leighton & M. Gierl (Eds.) *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge: Cambridge University Press. (pp. 61-84).
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1-4.

- Papert, S. (1980). Teaching children thinking. In R. Taylor (Ed.). *The Computer in School: Tutor, Tool, Tutee*. New York: Teachers College Press. (pp. 161-176). Retrieved from: <http://www.citejournal.org/vol5/iss3/seminal/article3.cfm>
- Papert, S. (2005). You can't think about thinking without thinking about thinking about something. *Contemporary Issues in Technology and Teacher Education* [Online serial], 5(3/4). Available: <http://www.citejournal.org/vol5/iss3/seminal/article4.cfm>
- Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, 139(1), 5-21.
- Reys, B. J. (1986). Teaching computational estimation: Concepts and strategies. In H. L. Shoen & W. J. Zweng (Eds.), *Estimation and mental computation – 1986 year book* (pp. 31-44). VA: National Council of Teachers of Mathematics.
- Reys, R. E., & Bestgen, B. J. (1981). Teaching and assessing computational estimation skills, *The Elementary School Journal*. 82(2), 116–127.
- Reys, R. E., Bestgen, B. J., Rybolt, J. F., & Wyatt, J. W., (1980). *Identification and characterization of computational estimation processes used by inschool pupils and out-of-school adults: Final report*. National Institute of Education (U.S.).
- Reys, R. E., Rybolt, J. F., Bestgen, B. J., & Wyatt, J. W. (1982). Processes Used by Good Computational Estimators. *Journal for Research in Mathematics Education*, 13(3), 183–201. Retrieved from <http://www.jstor.org/stable/748555>
- Richardson, A. R., & Waller, D. (2005). The effect of feedback training on distance estimation in virtual environments. *Applied Cognitive Psychology*, 19(8), 1089–1108. doi:10.1002/acp.1140
- Richardson, A. R., & Waller, D. (2005). The effect of feedback training on distance estimation in virtual environments. *Applied Cognitive Psychology*, 19(8), 1089–1108. doi:10.1002/acp.1140
- Rose, V. & Morgan, L. F. (1986). Teaching measurement estimation through simulations on the microcomputer. In H. L. Schoen & M. J. Zweng (Eds.). *Estimation and mental computation: 1986 yearbook* (pp. 204-211). Reston, VA: National Council of Teachers of Mathematics.
- Rubenstein, R. N. (1985). Computational estimation and related mathematical skills. *Journal for Research in Mathematics Education*, 16(2), 106–119.
- Samson, I. Z. (2004) U.S. Patent 6820800. Electronic Calculator. Retrieved from <http://www.google.com/patents/US6820800>
- Shiller, R. J. (2004). *Irrational exuberance*. New York, NY: Broadway Books. Accessed online at: <http://books.google.com>
- Schmidt, R, Young, D. E., Swinnen, S., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of experimental psychology. Learning, memory, and cognition*, 15(2), 352–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2522520>
- Shrager, J., & Siegler, R. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, 9(5), 405–410. Retrieved from <http://pss.sagepub.com/content/9/5/405.short>
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Instrument for Dissertation Study – Estimation Questions

- Siegler, R. S. & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, 75, 428-444.
- Siegler, R. S. & Booth, J. L. (2005). Development of numerical estimation: A review. In J. I. D. Campbell (Ed.). *Handbook of mathematical cognition*. (pp. 197-212). New York, NY, US: Psychology Press.
- Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation : Evidence for multiple representations of numerical quantity. *Psychological Science*, 14(3), 237–243.
- Schoen, H. L., & Zweng M. J. (Eds.). (1986). *Estimation and mental computation: 1986 yearbook*. Reston, VA: National Council of Teachers of Mathematics.
- Snow, R. E., & Lohman, D. F. (1993). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.) *Educational Measurement, Third Edition*. (pp. 263-331). Phoenix, AZ: Oryx Press.
- Sowder, J. T., & Wheeler, M. M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education*, 20(2), 130–146. Retrieved from <http://www.jstor.org/stable/749278>
- Stamper, J., & Koedinger, K. R. (2011). Human-machine student model discovery and improvement using data. In J. Kay, S. Bull, & G. Biswas (Eds.), *Proceedings of the 15th International Conference on Artificial Intelligence in Education* (pp. 353–360). Berlin: Springer. Retrieved from: <http://www.learnlab.org/research/wiki/images/8/86/Stamper-Koedinger-AIED2011.pdf>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
- Thomas, J. (2000). *A review of research on project-based learning*. The Buck Institute for Education: Novato, CA. Retrieved from: http://www.bie.org/index.php/site/RE/pbl_research/29
- Threadgill-Sowder, J. (1984). Computational estimation procedures of school children. *The Journal of Educational Research*, 77(6), 332–336. Retrieved from <http://www.jstor.org/stable/10.2307/27540072>
- Tillman, D. (2011). Performance assessment of digital fabrication activities with embedded mathematics pedagogy. In M. Koehler & P. Mishra (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2011* (pp. 898–901). Chesapeake, VA: AACE.
- Tillman, D., Ducamp, G., Dejaegher, C., Cohen, J., Kjellstrom, W. & Smith, S. (2012). A role for digital fabrication activities utilizing the engineering design process in developing preservice elementary teachers' mathematics and science pedagogy. In P. Resta (Ed.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2012* (pp. 1085-1090). Chesapeake, VA: AACE.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124– 1131.
- Usiskin, Z. (1986). Reasons for estimating. In H. L. Schoen & M. J. Zweng (Eds.). *Estimation and Mental Computation: 1986 Yearbook* (pp. 1-15).
- Wheatley, G. H. & Hersberger, J. (1986). A calculator estimation activity. In H. L. Schoen & M. J. Zweng (Eds.). *Estimation and mental computation: 1986 yearbook* (pp. 182-185). Reston, VA: National Council of Teachers of Mathematics.

Instrument for Dissertation Study – Estimation Questions

- WMS Factsheet. (2013). Demographic information for W. middle school. Retrieved March 3, 2013, from <http://www2.k12albemarle.org/school/WMS/about/Pages/fact-sheet.aspx>.
- Xie, Q., & Tinker, R. (2006). Molecular dynamics simulations of chemical reactions for use in education. *Journal of Chemical Education*, 83(1), 77–83.
- Yang, D.-C., & Hsu, C.-J. (2009). Teaching number sense for 6th graders in Taiwan. *International Electronic Journal of Mathematics Education*, 4(2), 92–109.

Appendix A: Problem Set for Study 4

Contents

Segment COMP-EX.....	124
Segment COMP: Straight Computation Estimation	124
Segment VOL: Application - Volume Estimation.....	124
Segment MORE-COMP:	128
Interview Questions	128

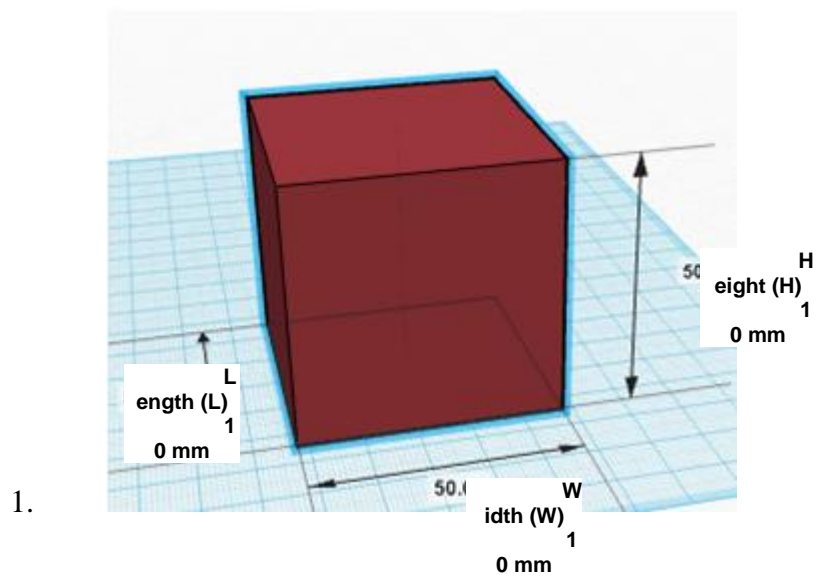
Segment COMP-EX

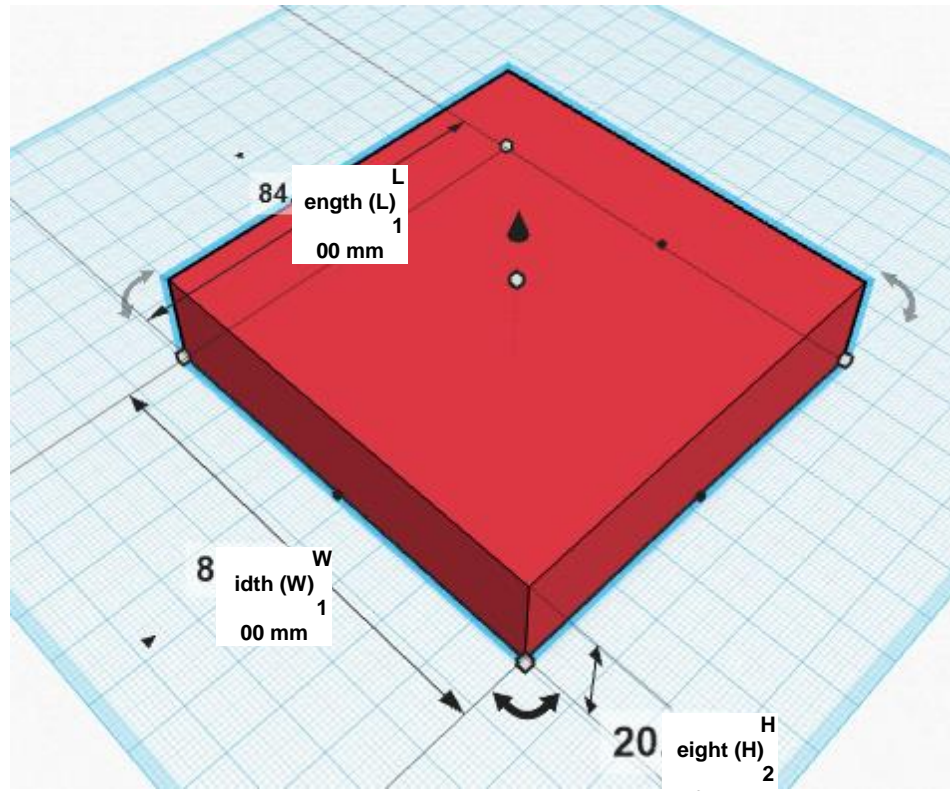
1. Round 17 to the nearest ten
2. Round 316 to the nearest hundred
3. Round 5676 to the nearest thousand
4. Round 2718 to the nearest hundred

Segment COMP: Straight Computation Estimation

1. 62×23 [from 2010 research]
2. $159 + 164$ [from 2011 research, $15,935 + 16,462$, simplified]
3. $491 - 417$ [from 2011 research, $49,167 - 41,705$, simplified]
4. $426 \div 7$ [from ACE, 1980, calculator segment, $42,163 \div 73$, simplified]

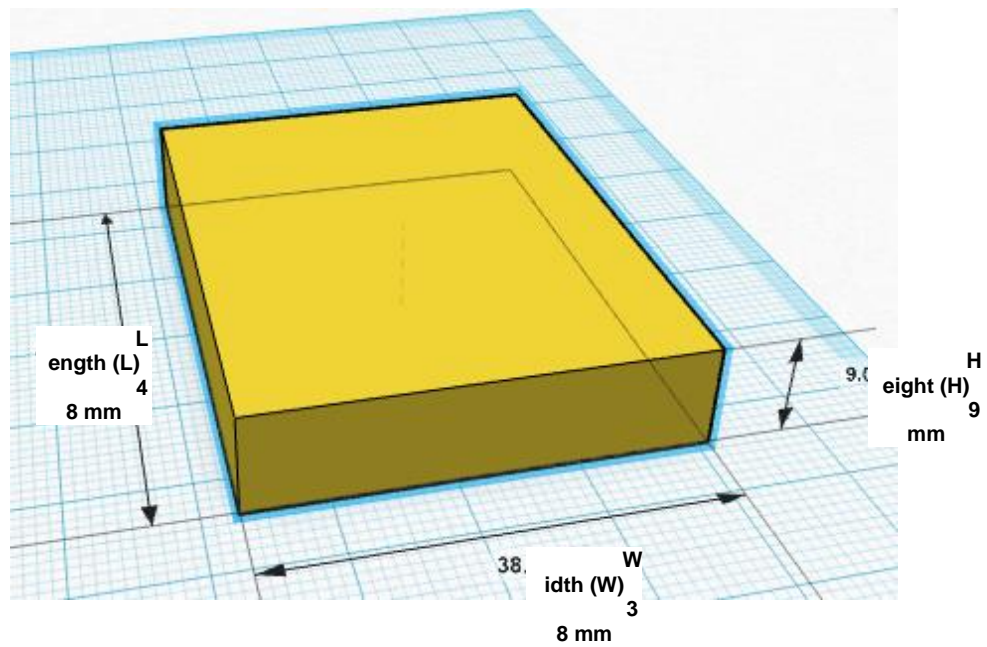
Segment VOL: Application - Volume Estimation



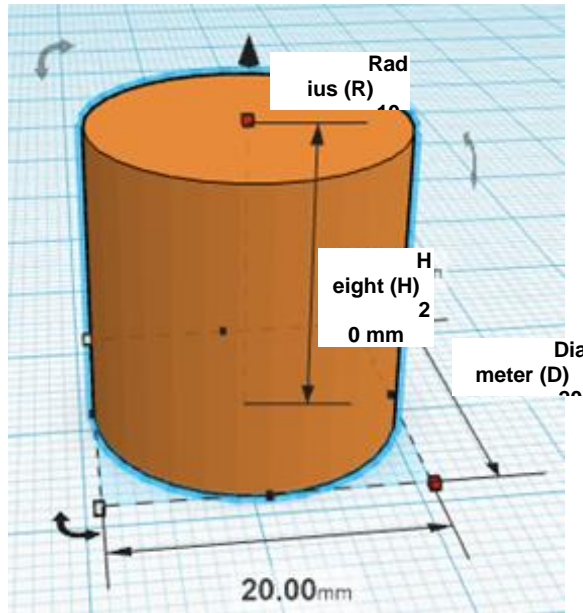


2.

estimate the volume of this shape [new for 2013]



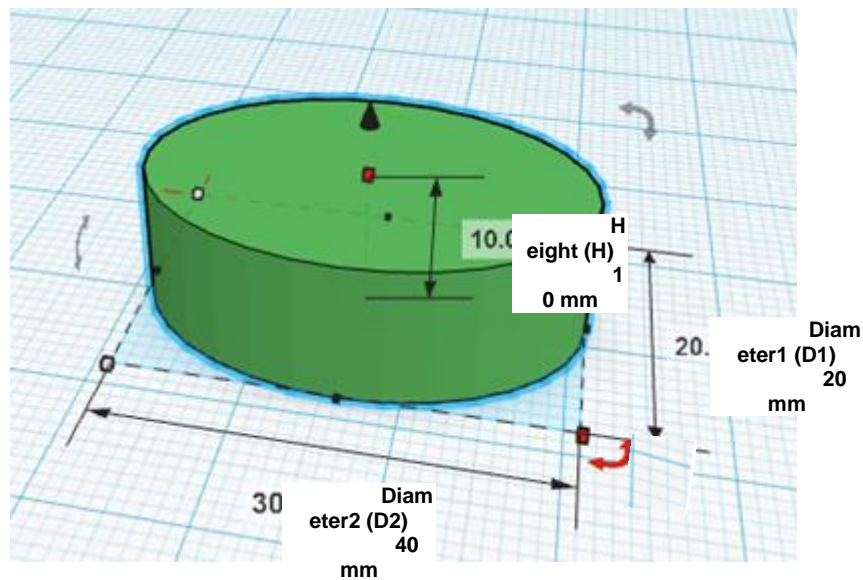
3.



4.

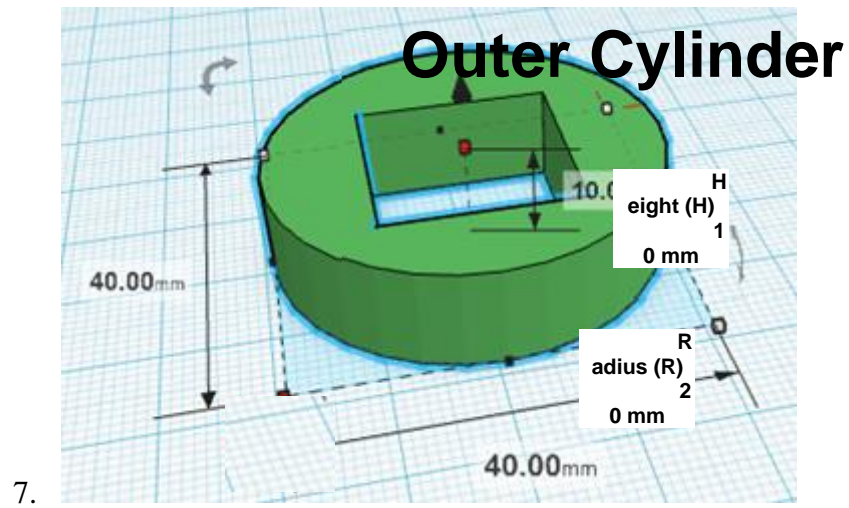
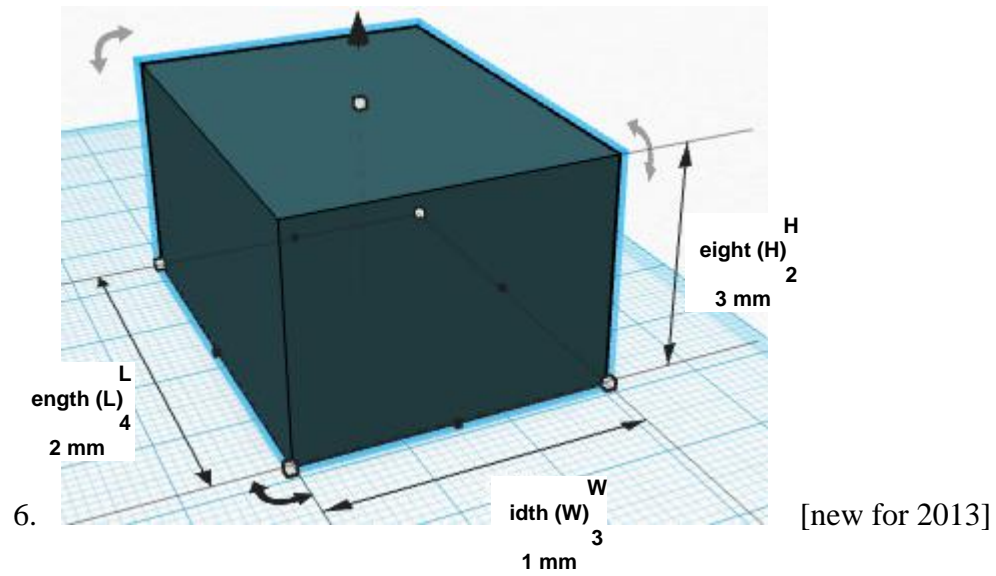
estimate the volume of this

shape [new for 2013]

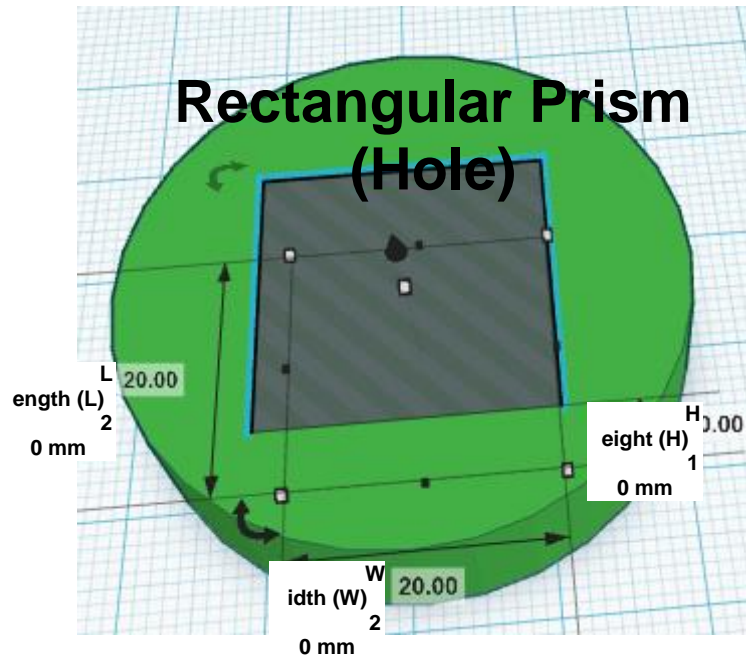


5.

estimate the volume of this shape [new for 2013]



1.



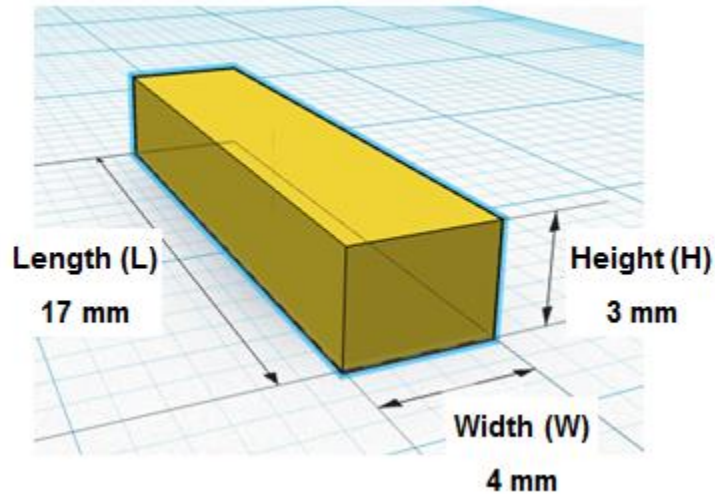
Segment MORE-COMP:

1. 87,419 [from ACE, 1980]
92,765
90,045
81,974
+ 98,102
2. $83420 - 56179$ [new for 2013]
3. $\frac{347 \times 6}{43}$ [from ACE, 1980]
4. $1\frac{7}{8} \times 1.19 \times 4$ [from ACE, 1980]
5. $474257 \div 8127$ [from ACE, 1980]
6. $67081 \times 35\%$ [new for 2013]

Interview Questions

1. $436 + 79$ [from ACE test, calculator segment: $436 + 972 + 79$, simplified]

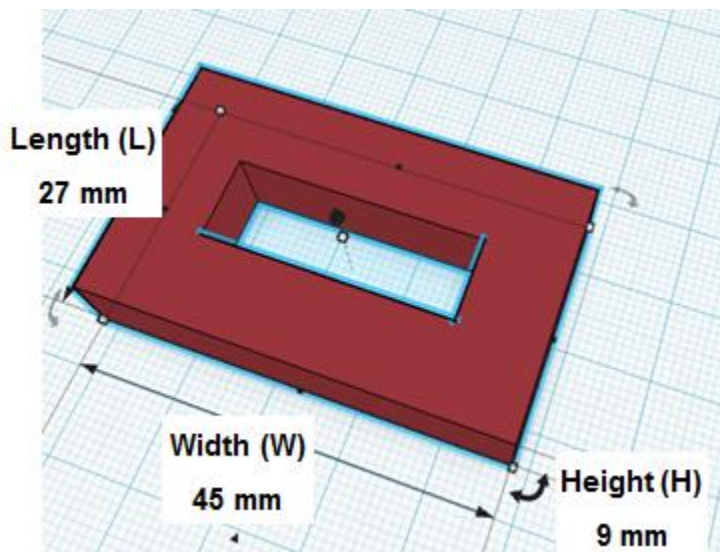
2. 896×19 [from ACE test, calculator segment]
3. $280 \times 50\%$ [new, for 2013 research]
4. $1000 \times 37\%$ [new, for 2013 research]
- 5.



Estimate the volume of this shape.

The volume of a rectangular prism is

$$V = L \times W \times H.$$



6. First estimate the volume of the entire shape.

Remember, the volume of a rectangular prism is

$$V = L \times W \times H.$$

Then estimate the volume minus the hole. The hole is 20% of the outside rectangular prism. (The rest of the shape is 80% of the volume.)

[new, for 2013 research]

Appendix B: Problem Set for the Interview Segment of Study 4

Name _____

7. $436 + 79$

8. 896×19

9. $280 \times 50\%$

Estimate the time needed to print the pencil holder in plastic:

10. (Observe the actual pencil holder). It is an outer cylinder minus an inner cylinder:

Outer: Radius (R) 40mm Height (H): 30 mm

Inner: Radius (R) 20mm Height (H): 30 mm

The formula for the volume of a cylinder is:

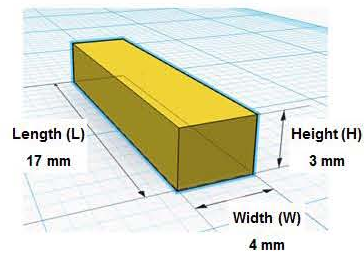
$$V = \pi \times R^2 \times H$$

You may use 3 as an estimate for π .

The time needed to print in plastic is:

$$\text{Time} = 2 \times V/1000 + 5$$

11.



12.

