The Effects of Cognitively Guided Instruction (CGI) on the Use of Representational Modeling by First and Second Grade English Language Learners (ELLs) During Individualized Assessments of Arithmetic Word Problem Solving.

Edwing A. Medina
University of Miami, edwingmedina@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
THE EFFECTS OF COGNITIVELY GUIDED INSTRUCTION (CGI) ON THE USE OF REPRESENTATIONAL MODELING BY FIRST AND SECOND GRADE ENGLISH LANGUAGE LEARNERS (ELLs) DURING INDIVIDUALIZED ASSESSMENTS OF ARITHMETIC WORD PROBLEM SOLVING

By

Edwing A. Medina

A DISSERTATION

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Coral Gables, Florida

August 2019
THE EFFECTS OF COGNITIVELY GUIDED INSTRUCTION (CGI) ON THE USE OF REPRESENTATIONAL MODELING BY FIRST AND SECOND GRADE ENGLISH LANGUAGE LEARNERS (ELLs) DURING INDIVIDUALIZED ASSESSMENTS OF ARITHMETIC WORD PROBLEM SOLVING

Edwing A. Medina

Approved:

Walter G. Secada, Ph.D.
Professor of Teaching and Learning  Luciana de Oliveira, Ph.D.
Professor of Teaching and Learning

Mary A. Avalos, Ph.D.
Research Associate Professor of Teaching and Learning  Ji Shen, Ph.D.
Associate Professor of Teaching and Learning

Cengiz Zopluoglu, Ph.D.
Assistant Professor of Educational and Psychological Studies  Guillermo J. Prado, Ph.D.
Dean of the Graduate School

Robert Schoen, Ph.D.
Research Faculty III, Florida Center for Research in Science, Technology, Engineering and Mathematics (FCR-STEM)
Florida State University
Cognitively Guided Instruction (CGI) is a professional development framework that informs participating teachers regarding research-based knowledge of elementary arithmetic-problem types and children’s developmental strategies for solving them. A classroom taught led by CGI participants includes one where heterogeneous learners (including English Language Learners, ELLs) are using and discussing representational systems and modeling tools (such as base-10 blocks, snap cubes, fingers, and verbal narratives) as part of the collaborative whole-class and small group problem solving process. This exploratory study investigated the effects of classroom instruction by CGI teachers on ELLs’ use of representational systems as they independently solved arithmetic word problems.

A subset from a large-scale randomized control trial, totaling fifty-two first graders and fifty-three second grade ELLs (including nineteen and twenty-five CGI ELLs, respectively), were observed and videotaped during an individualized, statistically-validated arithmetic assessment comprised of several question formats. Students’ use of different, total, and simultaneous representational
modeling tools were coded and tallied using video-coding software as up to five different word problem types were solved. The resulting data were analyzed using analyses of variance and effect sizes, comparing between subjects (Treatment versus Control) and within subjects (by Problem Type).

This study found that CGI ELLs were more likely to use a greater number of different and total representational modeling tools in more complex ways when solving a variety of word problem types, than their non-CGI peers. The first grade ELLs exhibited these findings more often and more consistently than their second grade peers. The significant differences in representational tool use by first grade CGI ELLs (when compared to their non-CGI peers) as they solve several word problem types emphasized in the curriculum (such as “Addition/Subtraction with a missing addend”; and “Multiplication with grouping”) appear to be reflecting CGI effects during assessments. Similar (though less pronounced and frequent) effects were found as second graders solved “Multiplication with Grouping” problems, an emergent 2nd grade curricular emphasis. Both grades showed CGI effects on representational modeling when solving “Subtraction with an unknown result” problem types.

This study’s findings are consistent with prior research in CGI that notes the mediating power of the word problem type. It contributes to that work by adding a representational modeling perspective as it suggests a transfer of the representational modeling practices present in CGI classrooms by ELLs when they are independently assessed on similar word problems and allowed to use similar representational modeling tools.
DEDICATION

I dedicate this dissertation to my family, whose love and support helped me start, persevere, and finish this endeavor. I honor my 83-year old Venezuelan grandmother, Hilda, who believed in me enough to know that her elementary school education and life experiences would still drive me to become the only person in our Venezuelan family to earn a doctorate. To my mother, Mabel, for whom words could never capture her love for and belief in me nor my gratitude to her. She was the first English Language Learner I ever met and continues being my greatest Cognitively Guided Instruction teacher. I’m glad I was finally able to give you a reason to drink that special glass of wine, mom! To my father, Udayram, who also persevered quite late in his life to earn his doctorate, he was my role model and my guide, constantly convincing me that I, too, could achieve this goal late in my life. To my sister, Shallini, her husband, Danny, and my nephews, Nathan and Gaven, who embody why and for whom we fight for educational equity, access, and accommodations, while trying to make every learning environment one that builds on what young people can do, their curiosity, and their current challenges. And finally, to my beloved Onyx, whose nightly purring and meowing while sitting on my reading materials and papers, helped me get through this process. Although you didn’t see me finish the journey, we finally did it, my Onyx.
ACKNOWLEDGEMENTS

I acknowledge the contributions to this effort by the members of the larger study’s Management Team: Drs. Robert Schoen (PI and Director of *Replicating the CGI experiment in diverse environments*), Mark Lavania, Amanda Tazaz (all from Florida State University), Juli Dixon (co-PI), and Kristopher Childs (both from University of Central Florida) for logistics surrounding the data that made this study possible.

The research reported in this study was funded by IES Grant Number R305A120781 (OMNI Number 186000-520-030852), Replicating the CGI experiment in diverse environments. The data gathering required for the larger study (upon which this work is derived from) was overseen by the Florida State University’s Office of the Vice President for Research – Human Subjects Committee from 2012 through 2018. The analyses of this study conducted for this dissertation were reviewed and overseen by the University of Miami Social-and-Behavioral-Sciences Institutional Review Board (FWA# 00002247) as HSRO Study Number 20171017 (from 2017 through 2019). No endorsements of the results and findings should be inferred from the funding and/or from the help that I acknowledge; and all errors and omissions are mine.

I would like to express my sincere gratitude to my committee members, Drs. Luciana de Oliveira, Mary Avalos, Cengiz Zopluoglu, and Ji Shen (all from
the University of Miami), and Dr. Robert Schoen. This dissertation, I hope, reflects the rich doctoral preparation they facilitated for me prior to my dissertation writing, as well as their insightful and critical feedback during and after the dissertation writing – all of which exponentially improved my study.

Finally, I can’t thank Dr. Walter Secada, enough for all his help, as my advisor and dissertation chair, in constantly assuring that I crossed the finish line with a product that reflected the years of guidance, support, and discourse that we engaged in while I was an admittedly “difficult” advisee.
# TABLE OF CONTENTS

## LIST OF TABLES

<table>
<thead>
<tr>
<th></th>
<th>ix</th>
</tr>
</thead>
</table>

## CHAPTER

### 1 INTRODUCTION

<table>
<thead>
<tr>
<th></th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representational Systems</td>
<td>2</td>
</tr>
<tr>
<td>Defining representational modeling</td>
<td>5</td>
</tr>
<tr>
<td>Strategies Used by First and Second Graders' to Solve various Arithmetic Problem Types</td>
<td>9</td>
</tr>
<tr>
<td>First and Second Graders' Use of Representational Systems in Solving Arithmetic Problems</td>
<td>14</td>
</tr>
<tr>
<td>Focus on Phases of Problem Solving</td>
<td>16</td>
</tr>
<tr>
<td>Cognitively Guided Instruction (CGI)</td>
<td>17</td>
</tr>
<tr>
<td>The CGI Framework</td>
<td>18</td>
</tr>
<tr>
<td>Formative Assessment</td>
<td>18</td>
</tr>
<tr>
<td>Interpretation of Representational Systems as Indicating Cognitive Functioning</td>
<td>20</td>
</tr>
<tr>
<td>Instructional Practices in the CGI Classroom: Focus on Explaining Representations</td>
<td>21</td>
</tr>
<tr>
<td>English Language Learners (ELLs)</td>
<td>23</td>
</tr>
<tr>
<td>Focus on the Interaction of Language and Representational Modeling for ELLs</td>
<td>25</td>
</tr>
<tr>
<td>CGI and ELLs</td>
<td>27</td>
</tr>
<tr>
<td>Gender, Representational Systems, and Mathematics Achievement</td>
<td>28</td>
</tr>
<tr>
<td>Testing and Assessment of Primary Grade Students</td>
<td>29</td>
</tr>
<tr>
<td>Primary Grades</td>
<td>29</td>
</tr>
<tr>
<td>Issues in Test Bias, Reliability, and Validity in Mathematics Standardized Tests</td>
<td>31</td>
</tr>
<tr>
<td>Cognitive Interviews and Think Alouds as Tools to Reveal Test Bias and to Increase Test Validity</td>
<td>33</td>
</tr>
<tr>
<td>Primary Graders’ Use of Representational Systems in Assessment</td>
<td>36</td>
</tr>
<tr>
<td>Differences Between In-Class and Assessment Settings</td>
<td>37</td>
</tr>
<tr>
<td>Research Questions</td>
<td>37</td>
</tr>
</tbody>
</table>

### 2 RELATED LITERATURE

<table>
<thead>
<tr>
<th></th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representational Systems and Tools as Instructional and Assessment Supports</td>
<td>39</td>
</tr>
<tr>
<td>Representational modeling on Assessments</td>
<td>43</td>
</tr>
<tr>
<td>Distinguishing Imagistic Representational Modeling Tools: Fingers vs. Base-10 Blocks &amp; Snap Cubes</td>
<td>46</td>
</tr>
<tr>
<td>Social/Collaborative Supports during Instruction and Assessments</td>
<td>48</td>
</tr>
<tr>
<td>Mathematics Register and Linguistic Challenges when ELLs Solve Word Problems</td>
<td>51</td>
</tr>
<tr>
<td>Focus on Phases of Problem-Solving</td>
<td>57</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Representational Tools Used Individually or Simultaneously with Others</td>
<td>60</td>
</tr>
<tr>
<td>Assessments, ELLs, and CGI</td>
<td>63</td>
</tr>
<tr>
<td>CGI Problem Types and Solution Strategies</td>
<td>65</td>
</tr>
<tr>
<td>Validity, Reliability, and Bias of Individualized and Whole-Group Assessments</td>
<td>68</td>
</tr>
<tr>
<td>Cognitive interviews</td>
<td>69</td>
</tr>
<tr>
<td>Students Enrolled in Mathematics Classrooms Taught by CGI and non-CGI Trained Teachers</td>
<td>70</td>
</tr>
<tr>
<td>English Language Learners (ELLs)</td>
<td>71</td>
</tr>
<tr>
<td>First and Second Graders</td>
<td>72</td>
</tr>
</tbody>
</table>

### 3 METHODOLOGY

| Study Participants | 74 |
| Assessment Instrument | 74 |
| Cognitive Interview Procedures | 80 |
| Coding Representational Modeling Tool Use During Cognitive Interviews | 83 |
| Quantitative Approaches to Analyzing Study Data | 87 |
| Why Quantitative/Frequency vs. Qualitative/Descriptive Narratives? | 89 |
| Statistical Methodology | 92 |

### 4 RESULTS

| Descriptive Statistics | 97 |
| Inferential Statistics | 98 |
| Research Question 1 - Overall | 100 |
| ANOVA: Grade 1 Variety of Representational Tools Used | 105 |
| Effect Size: Grade 1 Variety of Representational Tools Used | 107 |
| ANOVA: Grade 2 Variety of Representational Tools Used | 108 |
| Effect Size: Grade 2 Variety of Representational Tools Used | 108 |
| Research Question 2 - Overall | 109 |
| ANOVA: Grade 1 Total # of Representational Tools Used | 111 |
| ANOVA: Grade 2 Total # of Representational Tools Used | 112 |
| Research Question 3 - Overall | 114 |
| ANOVA: Grade 1 # of Double/Concurrent Representational Tools Used | 115 |
| Effect Size: Grade 1 # of Double/Concurrent Representational Tools Used | 115 |
| ANOVA: Grade 2 # of Double/Concurrent Representational Tools Used | 117 |
| Effect Size: Grade 2 # of Double/Concurrent Representational Tools Used | 117 |
| Focus on Correct Grade 1 Responses (Research Questions 1, 2, and 3) | 119 |
Focus on Correct Grade 2 Responses (Research Questions 1, 2, and 3)  
Summary of Results  

5 DISCUSSION  
Summary of Findings  
Conclusions  
   Relationship of Representational Tool Use to Problem Types  
   Effects of CGI Instruction during Individualized Assessments  
   Implications for the Assessment of ELLs  
   Implications for the Instruction of ELLs  
   Implications for CGI Professional Development  
Limitations  
   Possible Limits to CGI based on Differences Between Instructional and Assessment Settings  
   Limitations of the Study Design and Assessment Tool Design  
Recommendations for Future Research  

REFERENCES  

Appendix A – Representational Systems’ interactions with Counting  
Appendix B – Instance Codes and Labels – Definitions and Examples
LIST OF TABLES

Table 1. Various arithmetic problem types experienced by 1st and 2nd graders.  10
Table 2. Strategies used to Solve various arithmetic problem types.  12
Table 3. Number and gender demographics of all grade 1 and grade 2 treatment (CGI) and control (non-CGI) English Language Learners (ELLs) (and their teachers).  79
Table 4. Items (and types and numbers) in the Word Problem section  82
Table 5. Number and percentage of all grade 1 treatment (CGI) and control (non-CGI) respondents for each assessment question type.  99
Table 6. Number and percentage of all grade 2 treatment (CGI) and control (non-CGI) respondents for each assessment question type.  99
Table 7. Grade 1 Treatment (CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type  101
Table 8. Grade 1 Control (non-CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type  102
Table 9. Grade 2 Treatment (CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type  103
Table 10. Grade 2 Control (non-CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type  104
Table 11. Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Variety of Representational Tools Used by Grade 1 students.  106
Table 12. Variety of available representational tools used by grade 1 treatment (CGI) and control (non-CGI) respondents for each assessment question Type.  107
Table 13. Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Variety of Representational Tools Used by Grade 2 students.  108
Table 14. Variety of available representational tools used by grade 2 treatment (CGI) and control (non-CGI) respondents for each assessment question Type.  109
Table 15. Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 1 students.  112
Table 16. Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 2 students.  113
Table 17. Total frequency of use of concurrent (two) representational tools by grade 1 treatment (CGI) and control (non-CGI) respondents for each assessment question type.  116
Table 18. Total frequency of use of concurrent (two) representational tools by grade 2 treatment (CGI) and control (non-CGI) respondents for each assessment question type.

Table 19. Grade 1 Summary table of correct student respondents, p-value mean difference significance levels, and differences of averages of the variety, total, and concurrent (2) representational tool instances.

Table 20. Grade 2 Summary table of correct student respondents, p-value mean difference significance levels, and differences of averages of the variety, total, and concurrent (2) representational tool instances.

Table 21. Total frequency of use of concurrent (two) representational tools by grade 1 treatment (CGI) and control (non-CGI) respondents correctly answering each assessment question type.

Table 22. Total frequency of use of (two) concurrent representational tools by grade 2 treatment (CGI) and control (non-CGI) respondents when correctly answering each assessment question type.

Table 23. Summary of specific statistically significant Question Type and Treatment findings, by research question and grade.
CHAPTER 1

INTRODUCTION

The purpose of this study is to better understand how first and second grade students orchestrate visual representations when solving arithmetic word problems. Specifically, I contrast the use of representational systems (referred to as “representational modeling”) by English Language Learners\(^1\) (ELLs) enrolled in classrooms taught by teachers implementing a curricular framework against those used by students in classrooms taught by teachers following a different, more traditional curriculum when they correctly (or incorrectly) solve a small number of individually-administered arithmetic tasks. The variety, total, and simultaneous use of representational tools evident during video-recorded sessions of students’ problem solving is compared between ELLs whose mathematics teachers have had one year of professional development in the Cognitively Guided Instruction (CGI; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989) pedagogical framework and ELLs whose teachers did not partake in any CGI Professional Development activities). No comparisons are made between representational tool use of first and second graders interviewed and video recorded for this study, as the difference between the assessment questions asked invalidates such a comparison.

\(^1\) This study uses the term “English Language Learners” (or ELLs), as it is the term that provides the basis for how the students were selected by the study. In this case, the larger study used the school districts’ definition of the term used by the State of Florida, as detailed on the Florida Department of Education website: http://www.fldoe.org/academics/eng-language-learners/. Current practice amongst experts in the field is to use the term “Emergent Bilinguals.”
Representational Systems

The National Council of Teachers of Mathematics (2000) regards spatial sense as the use of spatial reasoning and visualization while solving mathematics problems. Wheatley (1991) further defined spatial sense in terms of imagery, which includes the construction, representation, and the dynamic transformation and manipulation of self-generated images. My working hypothesis is that the students in this study apply and develop their spatial sense through their visualizations of quantities and operations while creating or manipulating visual graphic representations. I investigate the use of these representations from a sense-making, solution-checking perspective, as conveyed during a cognitive interview that debriefed and clarified students' problem solving at the end of each problem.

By Representational Systems, I am referring to a subset-mix of four of the five mutually interacting, internal systems of representation proposed by Goldin and Kaput (Goldin & Kaput, 1996; Goldin, 1998): *imagistic*, *formal notational*, *planning/executive*, and *verbal/syntactic* control systems. These systems provide configurations that help “encode meaning for the person, and they can be externalized to communicate meaning to others” (DeBellis & Goldin, 2006, pp. 132-133). The *imagistic* system includes the visual/spatial and tactile/kinesthetic encoding evidenced in this study by students using their fingers, base-10 blocks, base-100 blocks, and other manipulatives.

---

2 “Visual-graphic representations” (VGR) and “modeling” are used somewhat interchangeably throughout this study because they generally indicate the presence and use of a provided manipulative (base-10 blocks or snap cubes), writing tool (markers), or one’s fingers evidencing one of two representational systems (*imagistic* or *formal notational*) observed in this study. These two terms differ slightly when focus is on the *planning and executive controls* representational system exhibited by students when choosing, sequencing, or combining the use of these manipulatives, writing tools, or fingers while solving math word problems.
and snap cubes while solving math problems. In line with prior studies on these representational modeling tools, students’ use of fingers during word problem solving were analyzed separately from their use of base-10 blocks and snap cubes. Furthermore, the base-10 cubes, rods, and flats (typically representing quantities of 1, 10, and 100 each, respectively) were analyzed separately allowing for a more nuanced analysis of how students represent the magnitude of quantities and of the arithmetic operations that comprise word problems than otherwise possible. The formal notational system includes the arithmetic notations and numeration systems evident when students used markers (or other writing utensils) provided to create arithmetic notations and numeration systems while solving math problems. The third system of representation called verbal/syntactic (Goldin & Kaput, 1996; Goldin, 1998) manifested itself through students’ audible utterances (narrations, responses, and/or explanations). The fourth inter-related representational system focused on students’ planning and executive controls, as evidenced (but not coded for in this current study) by the visible outcomes and sequencing of the heuristic and strategic decision-making engaged by participants when solving the arithmetic word problems during this study. Any collection or analysis of evidence of the fifth (affective) and final system of representation (Goldin & Kaput, 1996) was beyond the scope and focus of this study.

The NCTM standards note that “Representing ideas and connecting the representations to mathematics lies at the heart of understanding mathematics” (NCTM, 2000, p. 136). Many mathematics classrooms have visual-graphic
representations that decorate classrooms and that provide learning and teaching resources such as number lines across the wall above whiteboards, a ten-by-ten one-hundred count chart or multiplication array, colored chips, base-10 blocks, coins, 2- and 3-dimensional geometric shapes, current events charts and graphs summarizing data, and vocabulary-word walls. Having these representational modeling exemplars visible and readily available to students in learning settings is thought to make mathematics more accessible to ELLs and to augment the benefits of having English proficient (EP) students engage in authentic, frequent discussions with ELLs (Domínguez, 2011; Moschkovich, 2012). Since the combination of the representational modeling tools (fingers, base-10 blocks, snap cubes, and markers and paper) allowed/provided to students during the assessments analyzed for this study is typical of students’ everyday instruction experiences, their familiarity and use is likely to support students’ word problem solving.

Although visual-graphic representations are ubiquitous and they function as meaning-making tools in mathematics classrooms for many students and educators, their use as meaning-making tools by ELLs in assessment settings remains “largely unexplored” (Bustle, 2004, p. 417). The studies that have explored the use of these visual-graphic representations have typically done so during instructional learning settings, while students work in groups, under the guidance of a teacher (Fuson & Briars, 1990; Henningsen & Stein, 1997; Hiebert & Wearne, 1993; Stein, Engle, Smith, & Hughes, 2008). Furthermore, these studies investigate the use of representational systems over a limited time span,
while students are working on a particular task or goal (Niemi, 1996; Turner, Domínguez, Maldonado, & Empson, 2006) and they seldom focus on ELLs (Martiniello, 2008, 2009; Moreno, 2002). This study will contribute to prior work on representational modeling by exploring how these representational systems are activated during individualized mathematics assessments comprised of word problems when representational modeling tools are provided during problem solving.

Because representational models are often viewed as means to an end rather than as ends in themselves, they seldom become the focus of investigation and exploration in mathematics learning nor assessment settings. This lack of focus on representational models is inconsistent with findings that suggest that, to use these representations and visualizations effectively in classrooms and on assessments, students should be familiar with and experienced in the notations, sign systems, manipulatives, and representations. In this sense, representational models help make explicit mathematical objects and activities that would otherwise remain invisible as both an abstract and a mental construct (Dörfler, 2006). Under this perspective, therefore, mathematics is regarded as a practice that uses, produces, manipulates, and interprets signs and other representations.

**Defining representational modeling**

The visual graphics representations discussed throughout this study are part of a larger research base specified using many other terms. Among these
are: “inscriptions” (Cobb, 2002; Dörfler, 2006; Lehrer, Schauble, Carpenter, & Penner, 2000); “embodiments” (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Hostetter & Alibali, 2008; Núñez, 2006); “visualizations” (Giaquinto, 2007); “manipulatives” (Moyer, 2001); “representations” (Ainsworth, 1999 & 2006; Duval, 2006; Schnottz & Bannert, 2003), and “metarepresentational competence” (diSessa, 2004; Hammer, Sherin, & Kolpakowski, 1991; Sherin, 2000). This array of terms hints at the variety of parallel and related research exploring this topic over the past few decades.

The choice of which term to investigate was purposeful, aligned with my research goals, and was informed by prior work. General terms such as “representations” and “visualizations” and specific terms such as “embodiment” and “manipulatives” were avoided in order to be as inclusive as possible of representational modeling that went beyond what was seen or made visible (as in “visualizations”) and what might apply only when tangible, movable objects or the human body itself were involved (as in “manipulatives” or “embodiments”).

In consideration of the difficulty of mapping existing terms onto the design and goals of this study, I adapted and combined existing terms that would more effectively capture the analytical focus of this study: representational modeling. Although aspects of many of these terms and their similarities and differences have shaped this study, I chose to adopt Schnottz and Bannert’s (2003) notions of descriptive and depictive representations (p. 143), with Carpenter, Fennema, Franke, Levi, and Empson’s (1999) “(Direct) Modeling” (or “the child’s explicit physical representation of each quantity in a problem and the action or
relationship involving those quantities . . . ,” p. 22) to create the term

“Representational Modeling.” Schnotz and Bannert (2003) specifies descriptive representations as spoken or written texts, mathematical equations, and logical expressions that are comprised of symbols and signs that describe and relate objects to each other. They go on to distinguish depictive representations such as pictures, sculptures, or physical models as those that allow us to extract relational information [but] do not contain symbols for these relations . . . they possess specific inherent structural features that allow us to read off relational information, and they are associated with the content they represent through these common structural characteristics. (Schnotz and Bannert, 2003, p. 143)

The characteristics of the assessment used in this study (and the resulting representational modeling) analyzed for this study reflect these descriptions and resonate with Goldin et al.’s (1996 and Goldin, 1998) imagistic, formal notational, planning/executive control, and verbal/syntactic representation systems chosen for this study.

Representations lack a fixed referent which would otherwise give a representation an immutable meaning. Therefore, the meaning of each representation rests in its respective operations, symbols, and applications. Modeling “reflects the distinctions portrayed in the analysis of problem types” (Carpenter, Fennema, Franke, et al., 1999, p. 4), which will evolve from initial, intuitive strategies into more abstract ways of modeling a problem, such as counting strategies and number facts. This study, in part, seeks to observe and document the distinctions representationally modeled by students
in their consideration of text, symbols, operations, concepts, and applications in assessment problems.

One of several reasons why I chose the term Representational Modeling for this study is to account for the possible purposeful and narrated emulation of modeling activities and tool use involving representations during the assessment used in this study, that may resemble those modeled or facilitated by the CGI-trained teacher and fellow students. The enactments of these representational models may then be observed, characterized, and analyzed (by the assessment, teaching, and research community).

In this study, for example, evidence of the enactment of abstract counting of numerals may be seen through one or more representational modeling activities: fingers moving in specific kinesthetic patterns until a predetermined quantity has been represented; numerals written in sequences and cadence experienced before; the joining or separating of snap cubes or base-10 cubes, 10-rods, and/or 100-flats. Investigating representational modeling activities as either/both a purely algorithmic exercise, and/or as more “creative, inventive, explorative, and experimental” enactments (Dörfler, 2006, p.107), in turn, benefit from and support ongoing research and pedagogy. Appendix A details some of the prior studies’ findings regarding representational systems’ interactions with counting, which are taken as givens in this study.
Strategies Used by First and Second Graders’ to Solve various Arithmetic Problem Types

During the early-elementary grades, students begin to call upon a handful of strategies to solve a variety of much-studied arithmetic problem types. Table 1 provides examples of arithmetic problem types that may elicit modeling and representation practices from students: Join (Result, Change, or Start Unknown), Separate (Result, Change, or Start Unknown), Part-Part-Whole (Whole or Part Unknown), Compare (Difference, Compare Quantity, or Referent Unknown), Grouping, and Measurement or Partitive Division (Carpenter, Fennema, Franke, et al., 1999). The highlighted cells in the table reflect the problem types and provide general descriptions of the questions posed to students in the assessment used in this study.

Certain problem types may naturally elicit strategies that rely on the use of manipulatives or other visual representations (Carpenter & Moser, 1983). These strategies include counting-all, counting-on, decomposition, retrieval (Laski, Casey, Yu, Dulaney, Heyman, & Dearing, 2013), direct modeling, counting down, counting down to, and trial and error (Carpenter, Fennema, Franke, et al., 1999). These strategies are described in greater detail in Table 2, below. When students use them (the strategies) to solve problems, they also use visual graphic representations to support their reasoning.
Table 1. Various arithmetic problem types, descriptions, and examples experienced by 1st and 2nd graders.

<table>
<thead>
<tr>
<th>Problem Type</th>
<th>Problem Description</th>
<th>Problem examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Join</strong></td>
<td><strong>(Result, Change, Start Unknown)</strong></td>
<td><strong>Result Unknown:</strong> Miguel has 6 red cars. Maria brought 7 blue cars to Miguel's house to play with him. How many cars did Miguel and Maria have altogether?</td>
</tr>
<tr>
<td></td>
<td>“...involve a direct or implied action in which a set is increased [or decreased] by a particular amount... The action in the problem takes place over time: There is a starting quantity at Time 1...; a second (or change) quantity is joined [or separated] to [from] the initial quantity at Time 2...; the result is a final quantity at Time 3” (Carpenter, Fennema, Franke, et al., 1999, pp. 7-8)</td>
<td><strong>Change Unknown:</strong> Miguel had 6 red cars. Maria brought some blue cars to Miguel's house to play with him. They played with 13 total cars. How many blue cars did Maria bring to Miguel's house?</td>
</tr>
<tr>
<td></td>
<td><strong>Unknown Start:</strong> Miguel had some red cars. Maria brought 7 blue cars to Miguel's house to play with him. They played with 13 total cars. How many red cars did Miguel have?</td>
<td><strong>Unknown Result:</strong> Miguel has 9 red cars. Maria took 5 of the red cars home after playing at Miguel's house. How many red cars did Miguel still have after Maria left?</td>
</tr>
<tr>
<td></td>
<td><strong>Unknown Change:</strong> Miguel has 9 red cars. Maria took some of the red cars home after playing at Miguel's house. Miguel now has 4 red cars after Maria left. How many red cars did Maria take home?</td>
<td><strong>Unknown Start:</strong> Miguel has some red cars. Maria took 5 of these red cars home after playing at Miguel's house. Miguel now has 4 red cars after Maria left. How many red cars did Miguel have before Maria took her red cars home?</td>
</tr>
<tr>
<td><strong>Separate</strong></td>
<td><strong>(Result, Change, Start Unknown)</strong></td>
<td><strong>Unknown Result:</strong> Miguel has 9 red cars. Maria took 5 of the red cars home after playing at Miguel's house. How many red cars did Miguel still have after Maria left?</td>
</tr>
<tr>
<td></td>
<td><strong>Unknown Start:</strong> Miguel has some red cars. Maria took 5 of these red cars home after playing at Miguel's house. Miguel now has 4 red cars after Maria left. How many red cars did Miguel have before Maria took her red cars home?</td>
<td><strong>Unknown Change:</strong> Miguel has 9 red cars. Maria took some of the red cars home after playing at Miguel's house. Miguel now has 4 red cars after Maria left. How many red cars did Maria take home?</td>
</tr>
<tr>
<td><strong>Part-Part-Whole</strong></td>
<td><strong>(Whole or Part Unknown)</strong></td>
<td><strong>Whole Unknown:</strong> Miguel has 6 red cars and 7 blue cars. How many cars does he have? <strong>Part Unknown:</strong> Miguel has 13 cars: 6 are red and the rest are blue. How many of Miguel's cars are blue?</td>
</tr>
<tr>
<td></td>
<td>Represent no implied or direct action, nor a change over time, while focusing on the relationship of various sets. Either the original Whole quantity is unknown (and both/all of the parts are known), or a part is unknown (and the original and other/remaining part(s) is/are known).</td>
<td></td>
</tr>
</tbody>
</table>
| **Compare**  
| **(Difference unknown, Compare quantity unknown, or Referent unknown)** | Find differences (a third set) representing the amount that one set exceeds another, after comparing two disjointed sets | *Difference unknown:* Miguel has 13 cars. Sven has 7 cars. How many more cars does Miguel have than Sven?  
*Compare quantity unknown:* Sven has 7 cars. Miguel has 6 more cars than Sven. How many cars does Miguel have?  
*Referent unknown:* Miguel has 13 cars. He has 7 more cars than Sven. How many cars does Sven have? |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grouping</strong></td>
<td>Modeling the groups mentioned in the given problems (using tally marks, counters, blocks, or other representations) and then counting the total number of objects grouped and represented.</td>
<td><em>Given:</em> “Yiwen has 5 mango trees. There are 7 mangoes on each tree. How many mangoes are there all together?”</td>
</tr>
<tr>
<td><strong>Measurement Division</strong></td>
<td>Students directly model and then count (either first, or later) the number of sets called for in a problem. Students keep track of the number of skip counts until the problem’s given total, and then record the answer as the number of times the student skip counted or repeated the addition or subtraction.</td>
<td><em>Given:</em> “Lisa has 17 bracelets. She wants to put them in bags that hold 3 bracelets each. How many bags can she fill? How many bracelets are left over?” or “How many threes are there in 17?”</td>
</tr>
<tr>
<td><strong>Partitive Division</strong></td>
<td>Students directly model the objects or quantities of the problem into the correct number of groups either one at a time or in batches. They then count (either first, or later) the number of objects contained within the number of sets called for in a problem.</td>
<td><em>Given:</em> “Yiwen has 5 mango trees. There are the same number of mangoes on each tree. Altogether, there are 35 mangoes. How many mangoes are there on each mango tree?”</td>
</tr>
</tbody>
</table>
Table 2. *Strategies used to solve various arithmetic problem types.*

<table>
<thead>
<tr>
<th>Selected Problem-Solving Strategies</th>
<th>Defined as…</th>
<th>Sounds/Looks like… (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>counting-all</strong></td>
<td>counting each addend in an addition problem, and then counting the entire combination of addends</td>
<td>Ex. 2+3 = “One, two. One, two, three. One, two, three, four, five.”</td>
</tr>
<tr>
<td><strong>counting-on</strong></td>
<td>Begin a count at the next value higher than the first addend, and continue this count until the amount of the second addend has been represented.</td>
<td>Ex. 2+4 = “Three, four, five, six.”</td>
</tr>
<tr>
<td><strong>decomposition</strong></td>
<td>Separating, or decomposing, a given problem into simpler sub-problems. In the example, the problem is decomposed into one of doubles, and then returns the amount omitted when the doubles were composed.</td>
<td>Ex. 7+9 = “Seven plus seven is fourteen. Fourteen plus the two (from the nine) is sixteen.”</td>
</tr>
<tr>
<td><strong>retrieval</strong></td>
<td>Recalling the answer to the problem from a prior experience solving it.</td>
<td>“I remember that ____ plus _____ is _____.”</td>
</tr>
<tr>
<td><strong>direct modeling</strong></td>
<td>Using objects or fingers to represent given addends, subtrahends, minuends, sums, products, or other quantities given in a problem, and then counting the union of various sets.</td>
<td>When given: “Miguel has 13 cars: 6 are red and the rest are blue. How many of Miguel’s cars are blue?” A student: Selects a set of 13 cubes. Then selects 6 red cubes and lines these up right under the 13 original cubes. Student then starts placing blue cubes next to the red cubes (and under the 13 original cubes) until the two sets of cubes are the same size. The student then counts the number of blue cubes placed and reports this total as the answer.</td>
</tr>
<tr>
<td><strong>counting down</strong></td>
<td>A backward-counting sequence is initiated from a given, starting amount in a problem and continues for the given number of counts in the problem. The last number in the counting sequence is the answer.</td>
<td>When given: “Miguel had 13 cars. 4 cars rolled away. How many cars did Miguel still have to play with?” A student: Counts down from 13 by four numbers: “twelve, eleven, ten, nine.”</td>
</tr>
</tbody>
</table>
**counting down to**

A backward-counting sequence is initiated from a given, starting amount in a problem and continues until the given number in the problem is reached. The number of words in the counting sequence is the answer.

*When given:* "A bus had 13 children. Some got off on the next stop. Now there are 8 children on the bus. How many children got off the bus?"

*Student:* Counts down from 13 by four numbers: "twelve, eleven, ten, nine, eight." Student may hold up a finger for each word mentioned in the count: five fingers. This becomes the answer to the problem.

---

**trial and error**

For *Partitive Division* problems, since the number of objects in each group that would be needed to skip count by is the unknown being asked by the problem (Carpenter et al., 1999).

---

**skip counting**

Repeated addition, particularly by certain numbers such as three and five (Carpenter et al., 1999).

*Given:* "A restaurant puts 3 slices of tomatoes on each sandwich. How many sandwiches can they make with 21 slices of tomatoes?"

*Student:* Counts "3, 6, 9, 12, 15, 18, 21." With each count, the student extends one finger. When done counting, the student sees seven extended fingers and claims: "7. They make 7 sandwiches."
Children notice, emulate, and use strategies found in Table 2 in order to solve problems. They start by using physical objects such as fingers, base-10 blocks, and snap cubes; and then over time, they replace these directly-modeled aspects with more abstract counting strategies and number facts (Carpenter, Fennema, Franke, et al., 1999). From this perspective, therefore, problem solving provides a platform for students to learn (and for teachers to teach) about how symbols and models may represent problem situations (Carpenter, Fennema, Franke, et al., 1999). Students' use of symbols and direct modeling during assessment can be analyzed for information about problem difficulty or about errors during problem solving. Observations on how students represent and solve problems, thus, can be used to inform the teacher on possible pedagogical approaches to help students address any difficulties or errors.

First and Second Graders’ Use of Representational Systems in Solving Arithmetic Problems

When students have frequent opportunities to engage with modeling by using multiple representations as they solve word problems, they develop and provide access to their own thinking and reasoning (Kamii, Rummelsburg, & Kari, 2005; Kendrick & McKay, 2004). Students' generating and revising of publicly evident representational models reflect a dynamic interplay among internal, mental representations that characterize mathematics; or as Lesh, Landau, and Hamilton (1983) note: “to do mathematics is to create and manipulate structures… whether they are embedded in pictures, manipulative materials,
spoken language, or written symbols” (p. 268). Lesh et al. (1983) further claim that “the processes used in manipulating and creating these structures, comprise the ‘conceptual models’” (p. 268), or what I refer to as “representational modeling”, used by mathematics students to solve problems.

Students develop a greater awareness of word problems’ contexts and other relevant factors through their use of representational modeling (Greeno & Hall, 1997). Student use of representational modeling provide evidence of their comprehension of a story-based mathematical situation (Harries & Barmby, 2006; Hiebert & Carpenter, 1992). A student’s use of representational modeling tools to create representational models provides some insight into a student’s mathematical thinking (Fennell & Rowan, 2001; Kamii et al., 2005; Kato, Kamii, Ozaki, & Nagahiro, 2002; Outhred & Mitchelmore, 2000).

Prior work on representational modeling has been conducted within cognitive-assessment settings or during instruction. This study adds to that work through its focus on the context of word problems during formal testing. CGI teachers’ implementation of the CGI framework encourages their students to model problems by using concrete representations, engaging counting or modeling, or using representational schema of the relationships among and between the problem data and concepts (Carpenter & Moser, 1983). Carpenter (1985) notes that schema that represent key relationships play a central role in problem solving, with “Better problem solvers (having) more powerful and more flexible schemata” (p. 37). This study will provide evidence regarding whether
representational systems common in CGI instructional settings, are also evident in formal-assessment settings.

Student-generated representational models provide access to a student’s cognitive processes, affording teachers and other students the opportunity to discuss these representational models, and their language-, keyword-, and/or symbol-based problem-solving procedures (Harries & Barmby, 2006) as public events and processes. The cognitive interviews analyzed for this study will provide additional evidence about the utility of this data collection technique to access ELLs’ cognitive processes while solving word problems.

**Focus on Phases of Problem Solving**

Even the most elementary of mathematics problems require the problem solver to think, to recall, to (re)create, and to verify the steps that are taken from the moment the problem is posed until the moment it is solved and/or checked for correctness. This process supports the dissection of a mathematics problem solving process into a series of phases or steps that are taken in some form of linear, nested and layered, zig-zagged, or even circular fashion as the student perseveres towards the solution. Observed and annotated representational modeling and tools should be evident in different forms and at various frequencies throughout the problem-solving process. This study will consider the representational modeling that is evident between the moment a student is given
a mathematics word problem and the moment a final solution is declared (or when a student conveys a wish to stop attempting to solve a particular problem).

**Cognitively Guided Instruction (CGI)**

A key goal of Cognitively Guided Instruction (CGI) is to help students apply “some of the intuitive, analytic modeling skills exhibited by young children to analyze problem situations” so that they may “avoid some of their most glaring problem-solving errors” (Carpenter, Fennema, Franke, et al., 1999, p. 55). While there is a large body of research on how classroom processes help students to correct their errors during instruction, less is known about when and how students self-monitor, avoid and/or correct errors during individual assessments.

CGI supports the teaching of elementary school mathematics by encouraging and building on children’s problem-solving strategies including the use of representational tools. CGI teachers' professional development helps them to understand how students’ specific strategies and the representational tools that they use can be used to reveal their (the students’) reasoning when solving mathematics problems; and hence, it enables teachers to make instructional decisions based on students mathematical understanding of the problems they are solving.

CGI learning-settings both resemble and differ from assessment settings. Students may be seated in groups or encouraged to work individually for part of the time, with their teacher moderating the discussion about student answers and
the solution strategies that students used to derive these. Mathematics word problems are posed to students, often accompanied by a discussion about the problem’s context. Students are encouraged to use any representational tools that make sense to them, and then to explain their solution to a peer, a group, or the teacher. CGI Teachers then use students’ strategies and the representational modeling evident during problem solving to guide and to extend their students’ mathematics understanding of the current and subsequent problems which may be more or less challenging depending on the teacher’s instructional goals.

The CGI Framework

This study contributes to the extensive CGI research literature as it explicitly investigates the use of representational modeling by individual ELLs while solving word problems in a non-instructional setting.

Formative Assessment

A teacher’s assessment of what a student understands about a problem should inform the teacher’s decisions on subsequent instruction (Secada & Carey, 1990). Some formative assessments that include a focus on students’ explicit representations (such as symbols, diagrams, maps, pictures, and language itself) help make students’ thinking visible (Heritage & Niemi, 2006).
The formative assessment practice (of questioning students on the nature and rationale of their visible and/or verbalized solution processes) and the synthesis of students’ solution strategies are an integral part of the CGI professional development program. The assessment in CGI classrooms, therefore, is “formal as well as informal, written and oral, and done in group or individual sessions” (Fennema, Franke, Carpenter, & Carey, 1993).

Research on the assessment of English Learners has already noted that visual graphic representations should only be considered an effective accommodation for ELLs’ mathematics word problem solving if: (1) ELLs perform better with the visual representations than without them, and (2) the performance of EP students on mathematics word problems with and without visual graphic representations is comparable (Abedi, Hofstetter, & Lord, 2004). Abedi et. al.’s (2004) latter condition means that representational modeling operates solely on language factors that are hypothesized to be hindering ELLs’ success on these problems (Solano-Flores, 2011). This study provides additional insights into how ELLs in CGI and in non-CGI classrooms use representational models at different frequencies and in different combinations as compared to each other. Since I notated participants’ verbal utterances during problem solving (such as self-narration, self-questioning, and counting aloud), this study also researched how often ELLs verbalized technical, math-specific language while applying their everyday language to understanding mathematics (i.e., the mathematics register; Avalos, Medina, & Secada, W. G., 2015; Schleppegrell, 2007), although it did not discern these utterances to this level of granularity.
Interpretation of Representational Systems as Indicating Cognitive Functioning

Prior to CGI, research evidenced students' use and interpretation of representational systems to understand complex mathematics concepts and to transfer knowledge to new situations (Heritage & Niemi, 2006). Since students who see their teachers engaging in representational modeling may not necessarily mirror this practice in their own problem solving (Dufour-Janvier, Bednarz, & Belanger, 1987), teachers may need to be more explicit about the use and interpretation of representational systems when solving problems. Greeno and Hall (1997) suggest that practicing the interpretation and use of representational systems must be part of a social environment where students learn to “participate in the complex practices of communication and reasoning in which the representations are used” (pp. 361-362). The use of representational systems help students' initial proficiency with procedural knowledge resulting from greater exposure, before then developing knowledge in domains like multi-digit subtraction (Hiebert & Wearne, 1996). Several studies have also provided evidence of how the use and interpretation of representational systems may indicate cognitive challenges, such as the incorrect interpretation of an algebraic letter as the name of an object; for example, Kaput (1987) found that students interpreted the letter “t” to mean “tall girls”, so that 4t was incorrectly interpreted as “four tall girls”. Niemi (1996) provides an example of how open-ended assessment tasks (of fractions, in that study) that require students to integrate
and to interpret representational systems will elicit evidence of links that students make when justifying their solution strategies between (a) the problem as it is given and (b) the students’ own understandings of the problem’s underlying content and concepts.

**Instructional Practices in the CGI Classroom: Focus on Explaining Representations**

CGI teachers exhibit many instructional practices in their classrooms that focus on discussing representational systems. CGI emphasizes “shifts from teachers finding ways of representing mathematical knowledge for students to students constructing their own representations based on their intuitive problem-solving strategies” (Carpenter, Fennema, & Franke, 1996, p. 14). When younger elementary school students solve addition and subtraction word problems, they are seen directly representing the quantities, actions, and relationships stated in the problems; for instance, (i) objects are joined or leave a grouping or (ii) fingers are extended when added to a count or closed when subtracted; either (i) or (ii) are used to illustrate addition or subtraction problems, respectively (Carpenter et al., 1989).

By focusing small-group or whole-class discussion on the range of strategies that students come up with, a CGI teacher also focuses students’ attention on the “similarities between strategies using physical tools and those using written symbols” (Carpenter, Fennema, Fuson, Hiebert, Human, Murray, … Wearne, 1999, p. 45). With the passing of time and as the magnitude of the numbers increase, the physical, direct modeling strategies are “abstracted and
abbreviated as children naturally begin to use counting strategies and number facts" (Carpenter, Fennema, Fuson, et al., 1999, p. 49). This progression from smaller to larger number supports the development and discussion of more efficient and flexible symbolic representation systems that depend less on the base-10 manipulatives that students used in direct models. This study incorporated the progression from smaller to larger numbers when investigating the representational tool use of 1st and 2nd graders.

In preparing CGI teachers to anticipate these progressions while discussing representational systems observed in their students’ solution strategies, researchers have asked: “How should symbols be linked to the informal knowledge of addition and subtraction that children exhibit in their modeling and counting solutions of word problems” (Carpenter & Fennema, 1992, p. 461). Because students in CGI classrooms are frequently called on to explain their solution strategies, they need to, first, have reflected on these strategies, and then, to have planned their verbal explanations (Carpenter, Fennema, & Franke, 1996). These explanations begin as narrations of evident, completed procedures. With practice and time, however, the explanations students provide become the solutions themselves (Carpenter, Fennema, & Franke, 1996).
English Language Learners (ELLs)

Students who speak a language at home other than English are formally evaluated by school officials to determine their English proficiency. The designation of English Language Learner is a legal designation that generates funds and that entitles the student to receive special language support towards academic achievement in classrooms whose primary language of instruction is English (Driver & Powell, 2016). Teachers implement a number of pedagogical moves that would make their mathematics instruction more inclusive of ELLs including: integrating ELLs' primary language and home experiences into mathematics instruction, making available strategic support resources (such as gestures, representational tools, and visual graphic representations) that will enable ELLs to consistently and effectively participate in mathematics discussions, emphasizing conceptual reasoning and discourse over rote procedures, using contexts and native language familiar to ELLs to make mathematics content more accessible, and encouraging authentic, collaborative discourse amongst ELLs and EPs alike (Moschkovich, 2013; Taube & Jasper, 2009). This study evaluated the problem solving representational systems that were used by ELLs when some of the above-referenced, in-class accommodations that are typically used by teachers were not replicated (or may be far less explicit or evident) in the assessment settings.

Recent legislation mandates that older ELLs be assessed in large-scale, high-stakes assessments (Every Student Succeeds Act, 2015). These tests,
however, are written and administered in the language in which these students are still developing proficiency. Several recent efforts focused on investigating and mitigating the various language-based factors have been conducted to increase the validity of these written group-administered assessments. Two efforts found to be helpful in the construction and administration of valid assessments of ELLs are ensuring (1) the appropriate linguistic complexity of assessment items (Abedi & Lord, 2001; Martiniello, 2008; Secada, Medina, & Avalos, 2017; Shaftel, Belton-Kocher, Glasnapp, & Poggio, 2006) and (2) the appropriate use of academic language (Bailey & Butler, 2003; Butler, Lord, Stevens, Borrego, & Bailey, 2004; Guerrero, 2004; Hakuta, Butler, & Witt, 2000; Halliday, 1978; Scarcella, 2003; Stevens, Butler, & Castellon-Wellington, 2000). Both well-studied factors hinge on the language of the assessment questions. The English words comprising the assessment questions used in this study matter because they serve as the signified or represented objects that students focus on when engaging a representational system to help them solve a problem or check their work once a solution is found. This study investigated the representational systems used by students when problem solving, noting any patterns in the number of representational systems used by ELLs during an assessment, that are normally engaged in CGI and non-CGI classrooms. If this analysis can begin establishing patterns in the use of representational modeling tools, subsequent qualitative investigations may explore the characteristics of
ELLs’ representational models as well as any alignment of representational tool use to technical or everyday language that the arithmetic word problems they are solving are comprised of.

Focus on the Interaction of Language and Representational Modeling for ELLs

Despite mathematics’ long-held reputation as the “universal language” (West, 1933, p. 7), the complex, inter-connected, and constantly evolving nature of mathematics symbols, technical and natural language, and the visual, non-linguistic representations (O’Halloran, 2005) provide some sense of the difficulty most students have in learning to communicate their mathematics understanding in both classroom and assessment settings. Wilkinson (2015) argued that, when students consistently and successfully evidence the mathematics practices of communications and reasoning, they (the students) have become adept at using the mathematics register. This discipline-specific, language-dependent tool is comprised of highly- and semi-technical terms, dense noun phrases, math-specific and precise conjunctions, complex subordinated, and logical relationships (Celedon-Pattichis & Ramirez, 2012; O'Halloran, 2005; Schleppegrell, 2007). In assessment settings such as those analyzed in this study, the interaction of representational modeling and the mathematics register (both of which have the potential to help or hinder ELLs’ problem solving efforts) becomes critical because of the net effect that this interaction may have on student achievement. This study, while focused on patterns in the use of
representational models and tools, may help future researchers investigating the combined effects of language and nonlinguistic forms of representations on mathematics student achievement (Kachchaf, 2011; Kenney & de Oliveira, 2015; Martiniello, 2009; Solano-Flores, 2011; Solano-Flores, Barnett-Clarke, & Kachchaf, 2013).

Previous work by Fuson, Smith, and Lo Cicero (1997) relate the representational modeling choices evidenced when students of varied English proficiency counted and modeled counting. Fuson et al. specify three different concepts students engage while solving math problems in classrooms settings: unitary conception, decades-and-ones conception, and sequence-tens-and-ones conception. Unitary conception is when, for example, students count 34 cubes consecutively as “1, 2, 3, 4, 5,…32, 33, 34”. Decades-and-ones conception is evident when those 34 cubes are interpreted by a student to mean that the 3 represents the thirty decade and the 4 represents four. The sequence-tens-and-ones conception applies when a student counts the 34 cubes as “10, 20, 30, and then 1, 2, 3, 4. The findings represent one of the few research-based efforts at the nexus between representational modeling and language proficiency available in the related literature. This study expands this knowledge base, establishing relationships between representational tool use patterns and curricular problem types that elicited these, but not with their in-class experiences with representational modeling-rich problem solving. In turn, the findings of this study support making CGI’s implicit approach to representational modeling more
explicit. Furthermore, it supports the allocation of more time and space to facilitate students’ sharing of their representational tool-based problem solving approaches.

**CGI and ELLs**

The instructional characteristics of a CGI classroom are aligned with the characteristics of learning environments that support ELLs’ learning and achievement in mathematics classrooms (Celedón-Pattichis, Musanti, & Marshall, 2010; Musanti, Celedón-Pattichis, & Marshall, 2009). CGI emphasizes mathematics communication comprised of peer-to-peer, student-to-teacher, and student–to-class sharing, justification, and contextualized problem solving as recommended by Battey, Llamas-Flores, Burke, Guerra, Kang and Kim (2013) and Celedón-Pattichis, Musanti, and Marshall (2010), among others. CGI builds on the experiential knowledge that students bring to classrooms including using ELLs’ home language and informal understanding as resources as recommended by Battey et al. (2013), Celedón-Pattichis, Musanti, and Marshall (2010), and Musanti, Celedón-Pattichis, & Marshall (2009). CGI also makes use of students’ diverse strategies when solving problems, making meaning of technical mathematics and everyday language and concepts given in the problem, and working primarily on word problems (Battey et al., 2013). Alternatively, there may exist some learning settings where there may exist explicit or implicit beliefs regarding ELLs’ dampened intellectual ability, beliefs
that these students should participate in less challenging, less contextualized, language-free mathematics problems, and that these students should solve word problems using rote instructional techniques (Moschkovich, 1999, 2007, 2013;).

In CGI classrooms, however, the types of context and language-rich questions asked, and strategies and explanations encouraged and expected, characterize a classroom of high expectations, sophisticated thinking, and varied, concept-based solution strategies. CGI classrooms are led by a teacher who continues to learn how to support every student’s (including every ELL’s) efforts to succeed in the class (Battey et al., 2013). Even the variety of semantic structures that characterize the variety of word problems asked in CGI classrooms speak to mathematical aspects that vary the level of difficulty of the problems. These problem type variations make it possible for a greater variety of students to successfully participate in solving them (Secada, 1991, p. 214). CGI teachers, therefore, focus on what all their students know and understand, facilitate and encourage repeated opportunities for students to solve and explain a variety of word problems, and are prepared to help students extend that understanding.

**Gender, Representational Systems, and Mathematics Achievement**

This study acknowledges but does not account for gender differences in any of the analyses. The 1st and 2nd grade arithmetic word problems posed in this study do not reflect any of the advanced mathematics content (high school

---

3 Word problems involving actions are typically easier than those that do not; word problems with single sets are typically easier than those with two sets; problems where a whole set is the missing quantity are easier than those where the unknown quantity is a subset of a larger set.
geometry, trigonometry, and other rigorous courses) for which researchers have found differences in achievement and problem solving between genders (Battista, 1990; Burger & Shaughnessy, 1986; Linn & Petersen, 1985; Masters & Sanders, 1993). Furthermore, this study engages a sample of participants whose general mathematics achievement levels vary in line with typical groups of elementary school students, thereby minimizing the chance of finding any non-spurious gender differences. Given both of these considerations, I did not use gender (or prior achievement) as an independent variable in this study. Instead, the sample of video data drawn from the larger study’s data corpus has a balanced, diverse representation of ELL males and females representing the panoply of standardized assessment levels (prior to their participation in the assessment analyzed in this study) within the four dimensions of this study along grade in school (1st or 2nd grader) and treatment (enrolled in a classroom taught by CGI-trained teacher or not).

Testing and Assessment of Primary Grade Students

Primary Grades

The Florida Department of Education resembles many other states’ departments of education in its legislative charge to carry out the statutes that govern education policies, including when and how to assess students. These state-wide policies call for data on student academic-achievement and learning-
gains to be shared with numerous constituents: students, parents, teachers, school, administrators, and school district staff. Education researchers also use assessment data to inform and to gauge pedagogical and assessment initiatives and to evaluate national and international education comparison data. However, the statewide, standardized, comprehensive assessments in mathematics that are required by the State of Florida are administered annually and periodically between grades 3 through 10 (Fla. Stat. Ann., 2016). Therefore, the assessments that apply to the 1st and 2nd graders that participated in this study differ in their content, frequency, authorship, and stakes.

First and second graders participate in diagnostic, formative and summative assessments as well as in norm-referenced and criterion-referenced tests in their classrooms, even though these are not necessarily required, high-stakes, state-wide or national standardized tests (State of Florida Department of Education, 2009). These assessments may be comprised of constructed and extended constructed response questions, may include items enhanced by technology, may call for performance tasks, and/or may be ongoing, informal evaluations (Broward County Public Schools, 2017). Early-elementary school teachers constantly use diagnostic assessments in the classroom through observations, whole-class discussions, confidence indications, and informal dialogues with students. Formative assessments include exit slips, graphic organizers, self-assessments, and think-pair share activities. Summative assessments may include traditional portfolios and projects, as well as online apps such as Kahoot! and Quizlet (Nitko, 2001). All of these assessments
provide an opportunity for students to show what they understand and for
teachers to gauge where their instruction is effective and where it may need to be
delivered in alternate ways. Because these assessments include open-ended
questions similar to those used in this study, students should have been familiar
with the type of questions they were asked to answer prior to being assessed
using this study’s questions. Furthermore, the formative and diagnostic
assessments these 1st and 2nd graders are familiar with often include interactions
with and/or production of representational modeling tools and models.

Issues in Test Bias, Reliability, and Validity in Mathematics Standardized Tests

High-stakes tests and other types of formal assessment are administered
so that students can demonstrate what they have learned in class. These
assessments, designed to evaluate students’ mastery of the mathematical
practices, should be aligned to the Florida Mathematics Standards, which are
derived, in part, from the Common Core State Standards for Mathematical
Practice K-12 (National Governors Association Center for Best Practices &
Council of Chief State School Officers, 2010). The practices of relevance to this
study are: that students effectively use tools and build models when solving
problems, justify their solutions to others, understand and critique the
mathematical arguments of others, and make sense of their work (Wilkinson,
2015). Issues of equity, test reliability, and test validity become pertinent when
EP students and ELLs who are still acquiring English and who differ in their
levels of English proficiency (are required to take the same assessments and tests. Test bias and test validity remains as issues vis-à-vis the kinds of inferences that can be derived from test results when different groups of students use representational systems and tools\(^4\) in different ways; in other words, differential tool use may be factors that influence achievement. This study investigates how ELLs from different instructional settings (classrooms implementing the CGI framework and those following a traditional curriculum) use representational systems during individualized, word-problem based mathematics assessments.

Threats to an assessment's validity must be considered before understanding insights from assessment outcomes. When assessing ELLs, several validity threats are of concern (Abedi, 2004; Abedi, Hofstetter, & Lord, 2004; Abedi & Lord, 2001; Abedi, Lord, & Plummer, 1997; Celedon-Pattichis & Ramirez, 2012):

1. Language-based barriers when assessing mathematics word problems; this barrier may or may not be mitigated by ELLs’ use of provided or of student-generated representational modeling;
2. Whether an assessment values a product (such as a multiple-choice response) and/or a process (which may have had accurate, relevant visual graphic representations despite a potentially incorrect final answer) as the favorable or unfavorable evaluation of a student's mathematics word problem solution; and

\(^4\) such as base-10 blocks, snap cubes, and writing tools
(3) the similarity of or difference between assessment conditions and classroom conditions (given the absence of a teacher and fellow students that ELLs can interact with during assessments).

This study recognizes that mathematics word problems may be biased against ELLs, potentially affecting their cognitive understanding of these problems and ELLs' achievement outcomes resulting from these assessments. Experts recognize the inherent difficulty in distinguishing the mathematics content knowledge from the language skills tested via the word problems often found in assessments because English is the language used in most mathematics assessments (Abedi, 2004; Abedi & Lord, 2001).

Cognitive Interviews as Tools to Reveal Test Bias and to Increase Test Validity

The term “cognitive interview” conveys the notion that the person conducting the interview is following a more-or-less scripted protocol for observing and interviewing students as a means of gathering data about the reasoning that underlies performance on specific tasks that were administered during an assessment. This term, pertinent to this study, contrasts with “think alouds,” which convey the idea that students are asked to “think aloud” while doing each task and that scripted follow-up interview questions are seldom, if ever, asked. Representational modeling data from the types of cognitive interviews used in this study allow researchers and teachers to gather consistent
(or, in some cases, inconsistent) data as to whether specific mathematics problems tap into the mathematical concepts or underlying cognitive processes that test writers assume have been activated by specific items and/or tasks.

There are advantages and disadvantages to think alouds. Collecting data from short-term memory is preferable to relying on long-term memory because the short-term memory may be largely independent of interpretation by the student (Ericsson & Simon, 1993; Van Someren, Barnard, & Sandberg, 1994), particularly if the supporting data are evident generated representational models. Ericsson and Simon (1993) note, however, that data gathered during think-alouds may be problematic because they may be less coherent than interviews that may take place after the think alouds are completed. Another disadvantage to think alouds may arise if the cognitive load of problem solving and speaking (particularly for ELLs) is too great for the participant (Branch, 2000). Again, in such an instance, the focus on representational modeling could help alleviate the cognitive load a student might experience while participating in a think aloud that does not foreground representational modeling. This study incorporates findings from studies (Branch, 2000; Fonteyn, Kuipers, and Grobe, 1993) that showed asking participants questions after problem-solving provides useful data that made the real-time think-aloud data easier to interpret and understand. Therefore, this study reflected the use of cognitive interviews only, and not traditional “think alouds”. ELLs' unprompted verbal utterances such as self-narration or self-explanation while solving a problem are, of course, a form of
“think aloud;” and this study did make use of those data, but it did not explicitly prompt participants to engage in thinking aloud.

Cognitive interviews have been applied in several ways when working with ELLs. Most prior cognitive interviews (e.g., Jang, 2005) have been employed “to existing assessment data in order to determine structural information about which skills underlie good performance on the assessment” (Roussos, DiBello, Stout, Hartz, Henson, & Templin, 2007, p. 278). More closely aligned with the goals of the current study, however, some prior cognitive interviews have been specifically designed to diagnose problem solving skills. Cognitive interviews have also been suggested (Desimone & Le Floch, 2004) as a tool to develop more effective surveys because of the depth of understanding that they provide, particularly for individual items (rather than an aggregated data set).

Cognitive interviews also mitigate some of the three most common threats to the validity of data resulting from surveys (Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 1991). These three validity threats to surveys are: the complexity of phenomena of interest to researchers, the possibility that participants answer in a manner they believe to be the most socially desirable way, and the validity threat when an interviewer inadvertently provides misleading responses (Desimone & Le Floch, 2004).

Typically, ELLs participate in assessments without think alouds or cognitive interviews far more often than they do in assessments that include these additional data gathering methods (Nitko, 2001). Studies involving ELLs’ performance seldom provide (near) real-time evidence that the assessment
being used is tapping into the underlying mathematical concepts and cognitive processes as is assumed by the test writers and, implicitly, others using the tests. This study’s use of cognitive interviews during individualized assessments of arithmetic word problems affords researchers the opportunity to capture (and later analyze) students’ closest reflection of their mathematical thinking and contribute to the existing work in this area.

Cognitive interviews provide some insight into the nature of possible test biases that are not readily discerned from ELL students’ raw scores. This study uses cognitive interviews gathered via video data to focus on verbal utterances that resemble “think alouds” (but are not “think alouds” since these were unprompted and not always enacted) accompanying other representational tool uses during the problem solving phase of up to seven grade-specific word problems.

**Primary Graders’ Use of Representational Systems in Assessment**

Assessment items measuring data analysis, probabilities, and statistics domains tend to be more linguistically complex than word-problem items assessing algebra, relations, and patterns (Martiniello, 2009). Martiniello (2009) also found that algebra, relations, and patterns word-problems that were less linguistically complex also contained more schematic representational modeling (i.e., representations that include relatively-accurate spatial relations and proportions between objects and details that give salience to specific problem
components; Edens & Potter, 2007). Schematic representational modeling reduces the linguistic complexity of math word problems by helping ELLs make meaning of the texts that they read (Martiniello, 2009).

Differences Between In-Class and Assessment Settings

Are there assessment validity concerns when assessments are administered using protocols and settings that are acutely different from and do not reflect classroom practices and/or participants that used, created, and discussed representational models and modeling? Students in classes taught by teachers who promote and expect representational modeling to be a part of the problem-solving process (such as those familiar in CGI) may be at an unfair disadvantage in assessment settings that either ignore or discourage this type of representational practice as part of their mathematics problem solving.

Research Questions

This study answered the following research questions, which reflect the diversity of the participants and the quantitative nature of this study, and which builds on prior research in this area:

1/VARIETY) How many different representational modeling tools are used…

2/TOTAL) How many total representational modeling tool instances are used…

And
3/CONCURRENT) *How often are* representational modeling tools *used individually or simultaneously*…

…by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?
CHAPTER 2

RELATED LITERATURE

Representational Systems and Tools as Instructional and Assessment Supports

As mentioned in Chapter 1, this study uses the term “representational modeling” to refer to the constructs defined by four of the five mutually interacting, internal systems of representation outlined by Goldin and Kaput (Goldin & Kaput, 1996; Goldin, 1998): imagistic, formal notational, planning/executive control, and verbal/syntactic systems. Students’ use of their fingers, base-10 blocks, and snap cubes while solving the given mathematics word problems are three components of the imagistic system, as they help encode the visual/spatial and tactile/kinesthetic meanings for the problem solver and for the researchers. The arithmetic notations and numeration systems evident when students use the available markers or other writing utensils represent the formal notational system. Students’ heuristic and strategic decision-making, as well as the visible representational outcomes they created or used when solving math word problems reflected the third representational system of interest in this study: planning and executive controls. The fourth system of representation is comprised of audible discourse data during students’ cognitive interviews’ grammar, syntax, and natural language, distinguished as verbal self-narrations, responses, or other audible expressions. Only the independent word problem-solving phase of each student’s videorecorded assessment was analyzed for this study (sidestepping any analysis of the
Number Facts, Solving Equations, Equations: True/False and Multidigit Computation sections. Any representational tool uses of base-10 blocks, snap cubes, fingers, markers, or verbal explanations evident during the post-solution explanation phase of each word problem (if the interviewer pressed for the student to explain any part of the problem solving or solution after an answer was finalized) were not included in this study, even if they were evident.

Representational models have been used by students and teachers to demonstrate and to assess student understanding of mathematics content and concepts. Students and teachers may identify, introduce, or generate different static or dynamic, concrete or virtual, single or multiple representational models, both in classrooms and during assessments. Learners and educators may also use representational modeling tools and models to clarify, complement, or alternatively demonstrate their understanding of the mathematics in a problem they are trying to teach or solve.

Though much has been written about how students use representational models during instruction in CGI classrooms and about how their teachers and fellow students provide corrective feedback to help support the productive use of those models, little is known about how CGI-instructed students – especially English learners -- use those models within assessment setting. On the one hand, we might expect CGI-instructed students to have developed some familiarity and experience, both during and outside of formal classroom instruction, with representational modeling. We would expect students experiencing instruction that better acquaints them with recognizing the use of
relevant representational modeling tools, the need to generate these, the functional structures that comprise them, the practices necessary to carry out the relevant functions, and, possibly, the choices that may lead to errors or confusion (Dörfler, 2006), to transfer those skills into assessment settings thereby being able to carry out their problem-solving processes more successfully than students without that instruction. The familiarity that results from ongoing practice with representational modeling tools comes from developing and calling upon a cadre of recognized patterns, formulae, operations, relations, and their use that goes beyond rote memorization of symbols and visual graphic representations. An experienced user of these modeling tools should become adept at engaging representational modeling tools (from a repertoire of such representations) when solving tasks and questions like those where representational tools have proven effective when solving similar problems (or those that appear to be similar) in the past. The experienced user should also become familiar with the inferences supported by, as well as the limitations of, these tools.

Several studies have found that when solving addition, subtraction, multiplication, or division problems which were not already presented using representational models (such as the actions of, or numbers of, objects identical to the numerals in the problem), students use representational modeling tools and models to represent the relationships or actions prompted by the problem, referred to as direct modeling (Carpenter et al., 1989; Kouba, 1989). In a practice adopted by this study, these authors coded their observations as direct
modeling when they observed students using their fingers (or the provided representational modeling tools, such as snap or base-10 cubes) to directly model the content of a word problem. Unlike the current study where students are engaged in individual assessment settings, these prior studies documented students working in classroom learning settings either individually or in small groups under the guidance of their classroom teacher.

An analysis of their observations led to findings regarding the similarities between strategies that students displayed when using their fingers and/or the provided representational modeling tools (i.e., counters, plastic coins, base-10 blocks of various sizes and values). The Carpenter, Fennema, Fuson, et al. (1999) study informed some of the design of this study by investigating if the same strategies were used even when students used different representational modeling tools (p. 47). Because this study analyzes 1st and 2nd graders’ representational modeling tool use and models, Carpenter, Fennema, Fuson, et al.’s (1999) insights into the modeling differences of students in different grades and solving word problems with numbers of different magnitudes are relevant:

Over time, these physical modeling strategies are abstracted and abbreviated as children naturally begin to use counting strategies and number facts. Essentially the same pattern occurs for children’s solutions of problems with larger numbers. Children’s symbolic procedures evolve out of direct-modeling strategies with physical materials that incorporate groupings of ten. (p. 49)

This observation also alludes to the important developments in students’ understanding of the base-10 grouping and number system and place value.

However, the final design of this study established that the differences in the administered assessments’ concepts, numeric magnitude, and sequencing of
the problems challenged the validity and reliability of comparisons between 1st and 2nd graders’ representational tool use. As this study aims to establish a baseline of variety, frequency, and concurrent use of representational tools by ELLs while solving arithmetic word problems, any analysis of the use of representational tools vis-à-vis strategies and problem types across and between 1st and 2nd graders is beyond the scope of this study.

**Representational Modeling on Assessments**

Because representations are often used in mathematics assessments to execute calculations, they tend to become obsolete and forgotten as soon as the student shows proficiency and mastery with the abstract concept that the representations were originally called upon to assist. Mathematical thinking, at that point, returns to being a “mental manipulation of the mental objects (possibly supported by their external representations)” where “What is written or said is ([the] only?) expression of those mental processes,” (Dörrler, 2006, p. 101) and therefore the only aspect of a problem-solving process that can be evaluated for mastery and understanding.

Just as educators must make representation-based choices regarding whether/if to explicitly teach about representations or simply allow students to create their own, so too must authors of assessments make similar choices. Several studies found higher student achievement on problem-solving assessments when (a) assessment participants could independently create
representational models for the aspects of word problems that they (the students) regard as essential to solving the problem or (b) when students were explicitly taught how to use specific types of representations (Lewis, 1989; Van Essen & Hamaker, 1990; Wolters, 1983). Other studies have looked at the types of representations that students create on assessments. When schematic drawings (those resembling a diagram and depicting proportional spacing between objects) were compared to non-schematic drawings (which contain details unrelated to the problem’s solution), researchers found that students using the schematic drawings achieved greater math results on assessments than those who used less schematic visual representations (Edens & Potter, 2007).

Representational models and modeling, however, is not a panacea for all students nor even for many of their struggles with mathematics word problems. Prior studies have found that drawings provided to, taught to, or generated by students may result in incorrect solutions or an inaccurate, distracting representations of a key aspect of the word problem (Lean & Clements, 1981; Presmeg, 1986; Van Essen & Hamaker, 1990). Earlier findings (Moreno, Pirritano, Allred, Calvert, & Finch, 2006; Solano-Flores, 2011) showing that representations may be ineffective accommodations for ELLs in mathematics assessments also motivate some of the research questions of this current study. This study contributes to, but does not resolve, the seemingly contradictory findings that, in some cases, visual graphic representations may improve ELLs achievement on mathematics assessments while, in other studies, the opposite seems to be true, by investigating the conditions under which different
representational modeling tools are used and by whom they are used while solving a variety of mathematics word problems (Martiniello, 2009).

When assessments provide and prompt students to engage with and/or create representational models, as happens when word problems are used to assess mathematical problem solving, it seems natural to investigate the “visual grammar” of how these word problems are solved and to think of the representations as: “…(the) visual equivalent of adjectives (which modify nouns) and adverbs (which modify adjectives or verbs)” (Solano-Flores, 2011, p. 7). As with other types of assessments for which educators prepare their students, what students practice and learn during classroom instruction is assumed to be duplicated during assessments. Solano-Flores (2011) argues that the primary functions served by the most common visual illustrations during mathematics instruction consist of “metaphorizations (mainly, humor intended to make the students feel comfortable when they are taking tests) and reminding (e.g., visual clues intended to help students remember a discussion in class, or references to field trips and other experiences)” (p. 6). Other kinds of representational models can support student practice on additional problem-solving processes that, in turn, might help ELLs and EPs during assessments.

Mathematics assessments have incorporated the use of representational models for years. Increasingly, since the advent of high-stakes testing and the growth in the numbers of ELLs nationwide, these assessments have been tied to educational policies and funding that seek to improve the student achievement
and progress. In Florida, where this study took place, 9.2% of all the 2014-2015 public school students in the state were ELLs (McFarland et al., 2017).

The choice to use the *imagistic, formal notational, planning/executive control, and verbal/syntactic* systems of representation outlined by Goldin and Kaput (Goldin & Kaput, 1996; Goldin, 1998) was made because of their seamless mapping onto the design and goals of this research study. The study of *imagistic and formal notational* systems of representation afford the opportunity to quantify and analyze what representational tools students use when solving math problems; and the study of *planning/executive control and verbal/syntactic* systems of representation provide insights into how and when these representational tools were used. The selection of these constructs foregrounds the student as the doer of the *imagistic, formal notational, planning/executive control and/or verbal/syntactic* representational systems.

Students use representational systems regardless whether or not their teachers have had CGI professional development, the linguistic challenges that they may face, the type of problem being asked, and/or their grade in school, gender, or language proficiency.

**Distinguishing Imagistic Representational Modeling Tools: Fingers vs. Base-10 Blocks & Snap Cubes**

Although Goldin and Kaput (Goldin & Kaput, 1996; Goldin, 1998) provide a guiding 4-level (of 5 proposed) classification system for the mutually interacting, internal systems of representations analyzed in this study, the extant
literature supports further refining one of these categories into two distinct subcategories. While the imagistic system includes the visual/spatial and tactile/kinesthetic encoding evidenced by students using their fingers, base-10 blocks, and snap cubes while solving math problems, prior research has traditionally investigated these representational modeling tools separately (rather than together). Prior studies on the use of fingers in mathematical problem solving (Bender & Beller, 2011; Fuson, 1986; Fuson & Secada, 1986) seldom include investigations on base-10 blocks and snap cubes. The same holds true for studies on base-10 blocks and snap cubes (Falkner, Levi, & Carpenter, 1999; Carpenter, Fennema, & Franke, 1996; Fuson, Smith, & Lo Cicero, 1997; Chan, Au, & Tang, 2014) not including analyses of the use of fingers.

Because fingers often are used as an innate, initial representational modeling tool to then engage with other non-embodied representational modeling tools such as base-10 blocks and snap cubes, it makes sense to regard these two types of imagistic representational examples as separate phenomena. While there are few studies that investigate the use of multiple representational modeling tools by students during assessments, Flevares and Perry (2001) describe their work on teachers’ instructional use of multiple representational modeling tools (which they described separately as “Gesture”, “Picture”, “Object”, and “Writing”). The coding applied in the current study to document students’ use of fingers, base-10 blocks, and snap cubes (as well as their writing and verbal representational modeling tools), accordingly, separated the use of the imagistic tools of fingers from the use of base-10 blocks and snap cubes.
Regardless of age or culture, body parts including hands and fingers have often represented numbers and quantities (Butterworth, 1999). The sequencing of the fingers is one of the first signs of counting by a developing child, from a number of cultures (Ardila & Roselli, 2002). Verbal counting may actually follow finger counting in some children’s development (Brissiaud, 1992; Descoëtudes, 1921). Fingers help establish the one-to-one correspondence principle as they point to each object being counted in a set (Gelman & Gallistel, 1986; Gallistel & Gelman, 1992). Fingers can alleviate the burden on working memory (Alibali & DiRusso, 1999) by keeping track and increasing the accuracy of the items counted in mental calculations (Geary, 2005). Other cross-cultural studies have found that finger-counting habits have influenced the mental quantity-orientation related to number lines (Domahs, Moeller, Huber, Willmes, & Nuerk, 2010). The connection between counting and fingers is so explicit in some cultures that “digit” (from Latin’s digitus) in English or Spanish (dégito) means number and finger (Ardila & Roselli, 2002). While using our 10 fingers to count and represent numbers, this also helps to develop an understanding of the base-10 number system (Gracia-Bafalluy & Noel, 2008).

Social/Collaborative Supports during Instruction and Assessments

Fuson, Fraivillig, and Burghardt (1992) viewed “children as meaning makers who use conceptual structures to interpret what they see, hear, and feel” (Fuson, Smith, & Lo Cicero, 1997, p.739), while recognizing a Vygotskian view
of teaching (1934/1986) and a constructivist view of learning (Steffe & Gale, 1995). Conceptual structures are “hypothesized categories of quantitative activity that seem useful in understanding teaching and learning in a domain” (Fuson, Smith, & Lo Cicero, 1997, p. 740) based on prior research, on extensive experience with students solving problems in the domain(s), and on feedback from colleagues. Fuson et al.’s (1997) combined view calls on teachers to determine the conceptual structures students are engaging throughout “assessing-assisting cycles” (p. 739) that may last from only a few seconds (while helping a student in real-time) to much longer (when teachers methodically evaluate student work for signs of comprehension or misunderstanding). They call on teachers to contemplate “the conceptual tools that might help children understand the cultural words and written symbols of a domain (the mathematical referrers) as well as understand the quantities and operations of a domain (the mathematical referents)” (Fuson, Smith, & Lo Cicero, 1997, p.740).

CGI professional development acknowledges and integrates visual representations and manipulatives as a component of mathematical problem solving. Most of these learning and teaching approaches are carried out during in-class, whole-class and small group instruction (Solano-Flores, 2011). Few, if any, of these social, interpersonal supports from classroom experiences have analogs during individualized assessment settings. After the assessment is over, educators (and investigators) may use the completed exams to ascertain the strategies used by students when problem solving (including their selection of representational modeling tools, the representations they drew, and how their
representations map from the text of the problem onto their solution) (Jitendra, 2002). The data teachers collect from these analyses of student strategies may then inform future instruction. This instruction, however, is characterized by the kinds of representational modeling-based social and collaborative supports captured in the following CGI classroom exchange between the teacher (Ms. Elba; not her real name) and several of her ELLs, after posing the problem, “Ms. Sandra had six candies. Ms. Mary ate four. How many are left?”:

Ms. Elba: Muy bien. ¿Y qué pasó? ¿Ms. Mary qué? (Very good, and what happened? Ms. Mary what?)
Juan: Se los comió. (She ate them.)
Ms. Elba: Se comió … ¿Cuántos? (She ate… how many?)
Juan: Cuatro. (Four.)
Ms. Elba: O.K., ¿cómo le hizo? (How did she do it?) [Juan is staring at his drawing now very sure what to do next]
Cómaselos, miam, miam, miam [imitating the act of eating the candies] (Eat them… miam, miam, miam.)

As Juan is having trouble figuring out how to show the result of Ms. Mary eating the four candies, Ms. Elba involves the rest of the group, asking them:
Ms. Elba: ¿Cómo nos podríamos comer los dulces en un dibujo? (How could we eat the candies shown in a drawing?)
Student: We mixed it up [answering in English].
Ms. Elba: Oh, we mix it up [showing with her hand the act of erasing on the board].
Students: Lo borramos. (We erase them.)
Ms. Elba: Ah, lo borramos. (Ah, we erase them.) [Then, Juan is able to erase four candies. He counts how many are left and students conclude one more time that the result of Ms. Mary eating the four candies is that there are only two left for Ms. Sandra.]
(Musanti, & Celedón-Pattichis, 2013, pp. 52-53).

Through this exemplar exchange in a classroom reflective of the CGI framework, Ms. Elba helped scaffold her students’ solution progress by inviting them to engage in representational modeling by describing the mathematics concepts
they might have been thinking about using everyday language, after they drew the candies in the problem as circles. When the students shared that eating of the candies might be represented by "mixing up" of the candies, Ms. Elba gestured a sweeping hand movement that helped the students verbalize that eating the candies should be represented by erasing them from the drawing. Ms. Elba transformed a Separate (Result Unknown) problem that was initially Juan’s to solve, into an opportunity for social, collaborative supports engaging representational modeling. This, by definition, would not be features of an individualized mathematics assessment. While there is a feedback loop between instruction and formative assessments with regards to content, in general, and representational modeling in particular, there is no such parallel when students are in assessment-only settings.

Mathematics Register and Linguistic Challenges when ELLs Solve Word Problems

Research has noted teachers increasingly expecting their students to learn vocabulary, syntax, and the mathematical register, while constructing and communicating multiple meanings through sociocultural interactions with peers and their teachers as learning partners (Battey et al., 2013; Celedon-Pattichis & Ramirez, 2012). The role of vocabulary, ELLs’ cultural and linguistic experiences, and the very question of who is expected to generate and flesh out the meaning of words, contexts, and solution pathways of mathematical word
problems have been areas for which the need for more study has become evident.

Mathematical terms such as “rhombus” or “subtract” may be new to all students, but particularly ELLs, while other terms such as “sum”, “value”, or “product” may be familiar (and/or familiar-sounding) to ELLs while their grasp of the respective complex and specific mathematical uses of these terms may be limited at best (Freeman & Crawford, 2008). These complex and specific mathematical terms comprise some of the mathematics register.

A register is the collection of spoken words and written text whose use is mediated by the nature of the activity in which this language is operating (Halliday & Hasan, 1989). Romaine (1994) defines a register as “variation in language conditioned by uses rather than users, and involves consideration of the situation or context of use, the purpose, subject-matter, and content of the message, and the relationship between participants” (p. 20). Because registers tend to differ in semantics, they also have differences in grammar and vocabulary (Celedón-Pattichis, 1999; Celedon-Pattichis & Ramirez, 2012). The mathematics register includes, but is not limited to, the technical language generally regarded as mathematics. The mathematics register has developed to include “reinterpreting existing words in natural language such as ‘point’, ‘equal’, ‘reduce’, ‘(ir)rational’, ‘carry’, ‘column’, ‘set’, ‘power’, ‘table’, and ‘root’” to account for their diverse, evolving functions in mathematics and everyday discourse (Celedón-Pattichis, 1999, pp. 8-9). Analogously, when Cazden (1979) described mathematics as a language restricted “in both size and meaning” (p. 135), she
was noting how unlikely it was for ELLs and others to encounter most mathematical terms in non-mathematical contexts, so that any practice of English by ELLs with EP peers and adults will typically lack any practice including mathematics terminology, particularly in mathematical contexts. Finally, there are those mathematics register words that serve a similar function in everyday language, including ‘combine’, ‘and’, ‘increased by’, and ‘add’, whose similarity in both math and everyday contexts should be reinforced for ELLs by frequently using these terms in a variety of word problems (Celedón-Pattichis, 1999, p. 9).

Beyond single vocabulary words are the confounding effects of syntax, homophones, and mathematical symbols on ELLs in mathematics classrooms (Celedón-Pattichis, 1999; Freeman & Crawford, 2008; McLeman, 2012; Powell & Fuchs, 2010). Syntactic examples include comparatives such as “is less than”, “is greater than” or “x times as much” (Celedón-Pattichis, 1999, p. 9), and examples lacking a one-to-one correspondence such as “19/8” being read (in English) as “8 goes into 19” while its Spanish analog would be “dividir 19 entre 8” (or, literally, “divide 19 into 8”) (Castellanos, 1980). Another example of challenging syntax includes the concept of substitution, where we “substitute 5 for x” (so that wherever x was, 5 will now replace it), but when translated verbatim into Spanish (“sustituya 5 por x”), the meaning is reversed because now the instructions state to replace every instance of 5 with the variable x (Castellanos, 1980). An additional syntax example comes from a relatively well-known challenge to many students, including ELLs: symbolically misrepresenting “The number x is seven less than the number y” verbatim as “x = 7 - y” or “x = 7 <
y” instead of the correct “x = y – 7” or “x + 7 = y” (Celedón-Pattichis, 1999).

Syntactic variations, which present difficulties to many students, also challenge ELLs: “‘How many were there in all?’, ‘How many were there then?’, ‘How many more were there?’, ‘How many were left?’, and ‘How many fewer were there?’” (Celedón-Pattichis, 1999, p. 11; Ron, 1999). Logical connectives such as “but”, “either…or”, “if…then”, and “suppose” not only may cause ELLs difficulties in their new language, but students may also lack the proficiency to understand these in their home language as well (Dawe, 1983). ELLs may confuse homophones such as “than” with “then” and translate it into Spanish incorrectly (as “entonces”), or “many” with “money”, “took” as “talk”, “know” and “now”, and “exist” with “exists” (Celedón-Pattichis, 1999, p. 19), all of which add a linguistic challenge that derails an otherwise productive mathematical solution pathway.

As with the mathematics register, misunderstandings surrounding these and other syntactic, homophonic, and symbolic elements may reveal themselves in incorrect representational models generated by ELLs when solving word problems that include these elements.

When considering register entries that have multiple different (or similar) functions in both mathematics and everyday language, mathematics terminology that has no everyday language parallels, vocabulary, syntax, and general discourse, it is easier to understand the multiple sources of difficulties faced by ELLs learning mathematics and English. While researchers (Cardelle-Elawar, 1995; Celedon-Pattichis & Ramirez, 2012) encourage the use of students' home language when attempting to understand a mathematics problem, this presumes
a mathematics proficiency in that language by the students. Secada (1991) found that ELLs whose experiences do not resemble the vocabulary, language, or syntax in the word problems are likely to encounter difficulties. Mathematics instruction that remains at the basic word level with a focus on vocabulary translation misses the opportunity to discuss the more complex mathematics register of words, phrases, and ideas that are different from their everyday language parallels.

Secada (1991) notes that mathematics word problems are composed of “linguistic artifacts… logical, semantic, and syntactic features that children must interpret and understand” (p. 214) before they can successfully solve these problems. These language features, in turn, help indicate whether an action on the part of the problem solver must take place, whether the problem calls upon the use of single (or more) sets of items, and whether the answer will involve finding an entire set of items or a subset of them (Secada, 1991, p. 214). In addition to the nature of these features affecting problem difficulty, these features will also play out in representational modeling efforts as word problems that involve action, and/or a single set of items, and/or finding a whole (rather than a part of a) set have been shown to be easier to solve than other types of problems (Carpenter & Moser, 1983; Riley & Greeno, 1988; Secada, 1991).

The heterogeneous group of students that participate in CGI classrooms and partake in their assessments are both EPs and ELLs. Many of the ELLs show evidence of using representational modeling when solving word problems on their own during mathematics assessments. ELLs make use of nonlinguistic
visual representations (such as diagrams and symbols) as important sense-making tools when conveying understanding of mathematics concepts and solving tasks, particularly when their English proficiency limits how they would otherwise communicate their mathematical ideas (Barton & Neville-Barton, 2003, 2004; Turner, Dominguez, Maldonado, & Empson, 2006).

Nonlinguistic schematic representations (those closely resembling a diagram having proportional spacing between objects) could help mitigate the potentially negative effects of linguistic complexity of mathematics word problems by helping ELLs make meaning of the text in the problems (de Oliveira, 2012; Martiniello, 2009). When ELLs take math assessments that include different modalities of representation, ELLs’ symbolic, graphical, and representational understanding may help these students overcome their difficulty in comprehending and solving the problems (Barton & Neville-Barton, 2004).

Although students, including ELLs, enter school with some competence in solving addition and subtraction word problems (Secada, 1991), predominantly English text-based teachers’ instructions will call upon students to note key words and disambiguate terms that have multiple meanings to discern the relevant meaning to a problem’s text or context. Nonlinguistic representational models present in the original mathematics problem need to be interpreted and decoded for their meaning. These student-decoded representations become a part of the representational modeling tools engaged when problem solving mathematics scenario assessments.
ELLs may substitute the complex mathematics register with symbolic and graphic visual representations, using these to evidence as many of the mathematics practices as they can that don’t hinge on their English proficiency. Students engage these representations to support their understanding of emerging and well-understood mathematics concepts (Heritage & Niemi, 2006). Effectively doing so helps ELLs and EP students understand the representational models that may be provided when posed as a question and/or that students may generate while solving the word problem, as well as the mathematics problem’s everyday language and mathematics-specific terminology (Abedi, Lord, & Plummer, 1997; Halliday, 1978; Lemke, 2003; Martiniello, 2009; Mestre, 1988). This study provides some data to support teachers and test-makers continuing to incorporate the mathematics register as it relates to the representational models and tools that students are asked to use.

**Focus on Phases of Problem-Solving**

Mathematics problem solving in general, and word problem solving in particular, are seldom one-step start-finish events. Instead, they are typically thought of, executed, and discussed as a series of not-necessarily-unique steps that must be undertaken in order to reach a correct response. Many researchers have described these steps and phases in even more varied ways over the years, both explicitly and implicitly discussing the steps to solve a problem, which could include preparatory steps prior to launching into solving a problem,
operators and circumstances that mediate the transition from one step or phase to another, as well as steps that might be conducted after a solution is offered.

Greeno (1978) describes mathematics problems as segmented transformation problems comprising of an initial stage, a goal state, and problem-solving operators intended to reduce differences between each phase of a problem’s solution process and its end goal. Others have suggested that novices (but not experts) use means-end analyses of mathematics problems (similar to those used when solving physics problems) to work backwards from a problem’s objective (of solving for an unknown answer) through sub-goals to the initial, given aspects of the problem containing no unknowns, before reversing this procedure towards a solution through a forward-working sequence (Larkin, McDermott, Simon, & Simon, 1980; Simon & Simon, 1978). These authors note that expert problem solvers are able to work forward immediately without the need to backtrack from the end goal.

Garofalo and Lester (1985) used metacognitive behaviors to segment their four-phase cognitive/metacognitive mathematical task solving framework: orientation, organization, execution, and verification. As the name of the first phase implies, the student finds him/herself trying to understand and assess the mathematical task at hand in the orientation phase by engaging strategies that will align familiarity of this problem to prior experiences, an analysis of the problem’s given factors and potential struggle points, and, most relevant to this study, initial efforts at representational modeling and tool use. The organization phase leads students to identify local and broader goals and subgoals as a
solution is sought. This then sets up the carrying out, or the execution phase, where the student carries out goals set during the organization phase, and then monitors the actions taken to effect goals and subgoals, weighing these against factors such as accuracy or speed. The final phase calls for students to verify the decisions and outcomes of the goals and subgoals assessed, planned, and executed in the prior 3 phases. What authors like Garofalo and Lester often sidestep in their otherwise-meticulous dissection of the problem-solving process into phases, is the explicit role that representational modeling and tools play throughout the problem solving process. When representational modeling tools are explicitly mentioned, they are included for their function in support of the phases (as an “external memory source such as pencil and paper” when needing to remember subgoals, for example; Sweller, 1988, p. 273), rather than as its own phase(s) in the problem-solving process.

As with many enumerated things, there is not a general agreement on the number of steps a mathematics problem solving process should be, let alone what those steps are comprised of. In addition to Greeno’s (1978) two steps noted above, and Garofalo and Lester’s (1985) 4 phases, also noted above, others have proposed other segmentations of this process. Schoenfeld (1985) has segmented the problem-solving process into three phases (Analysis, Exploration, and Verification), while Verschaffel et al. (1999) specify five phases to this process (1. Build a mental representation of the problem; 2. Decide how to solve the problem; 3. Execute the necessary calculations; 4. Interpret the outcome and formulate an answer; 5. Evaluate the solution).
During my graduate assistantship as a doctoral student, I participated in a study with Dr. Jennifer Krawec at the University of Miami that built on Marjorie Montague’s (2008) work on the *Solve It!* Project. At the core of this study, was a 7-phase process (Read, Paraphrase, Visualize, Hypothesize, Estimate, Compute, and Check) with 3 embedded directives in each (“Say, Ask, Check”) that was essentially a restatement and reorganization of Polya’s (1957) heuristics. These phases and sub-steps explicitly and repeatedly asked students with learning challenges to engage representational modeling and tools throughout the problem-solving process. As a component of the Check phase, for example, students are explicitly instructed to “Say ‘Make a drawing or a diagram’”, “Ask ‘Does the picture fit the problem?’”, and “Check the picture against the problem information” (Montague, 2008, p. 41).

Notwithstanding the many ways in which the problem-solving process can be analyzed, this study conducts a similar representations-based analysis of problem solving within the context of assessment.

**Representational Tools Used Individually or Simultaneously with Others**

Ainsworth (2006) argues that, “Unfortunately, it is not sufficient to consider each type of representation in isolation -- representations interact with one another in a form of ‘representational chemistry’” (p. 185). When students engage multiple representational tools and systems, they may enjoy a “flexibility in the way that information is distributed over representations, impacting both
upon the complexity of each representation and the redundancy of information between representations” (Ainsworth, 2006, p. 185). The use of multiple, concurrent, non-redundant representations may harness the unique, different content that each of the representational tools may convey. Ainsworth (2006) also describes the simultaneous use of multiple representational tools as possible cases wherein the tools are (a) fully redundant so that each tool is conveying the same amount and granularity of information; or (b) partially redundant so that some of the information conveyed by each of the representational tools mirrors information conveyed in another tool and so that some information is unique to a particular representational tool. When there is little or no redundancy of content conveyed by a combination of representational tools used, students may then need to call upon and integrate additional representational tools that bridge, overlap, and supplement the current tools being used (Ainsworth, 2006). For example, a student might use fingers to count the Base-10 rods that they are using, and yet benefit from a verbalization of a counting by tens while the student touches the Base-10 rods, in order to convey a quantity that may not have been evident had any one of these three tools been used in isolation. It quickly becomes evident that coordinating the affordances conveyed by each representational tool, along with how these reinforce, supplement, or contradict each other when used simultaneously, has the potential to benefit as well as to adversely affect learners.

A number of studies have found mixed results when it comes to the usefulness of specific combinations of representational tools used, such as
whether spoken and written text presented simultaneously are beneficial (Kalyuga, 2000; Mayer & Sims, 1994). Additionally, researchers have found that the translation between (Tabachneck, Leonardo, & Simon, 1994) and the integration of (Kozma, Chin, Russel, & Marx, 2000) multiple, simultaneous representational tools for a novice vs. an expert student may also affect on whether the use of concurrent representations is a benefit or hindrance.

Ainsworth (2006) suggests that one of the reasons why the use of multiple representational tools may result in different benefits is because the multiple representational tools may play different pedagogical and assessment functions. If, for example, multiple representational tools are used to support a variety of computational properties but a student doesn’t need more than one of the representational tools, the addition of a second or third representational tool may add nothing to the student’s learning; and it could detract from their problem solving progress. On the other hand, if the use of multiple representational tools helps one tool constrain the other or to focus the student’s attention on a particular aspect of the representational tool’s affordances, and/or if both representational tools supplement each other’s information conveyed, then the concurrent use of representational tools may be beneficial to the student.

Because students still need to decide at what point to add a new representational tool or switch between them, there is potential for a beneficial combination of representational tools to be used, either by luck, or as a result of guided and/or independent practice. If the use of multiple representations is possible, then students need not be limited by a particular representation’s
weaknesses or its strengths. In order for students to reap the benefits afforded by using multiple representational tools, they must understand how different representations relate to and operate on each other (Ainsworth, Bibby, & Wood, 2002). This type of synthesis and understanding may develop through practice in classrooms where CGI is practiced, as the realizations of the different affordances of singular and multiple tools used can come from students as much as from their teacher.

Assessments, ELLs, and CGI

A word-based scenario problem is “a mathematics calculation embedded within sentences” whose (typically) English text is used to “identify missing information, make a plan to solve the problem, and perform one or more calculations to get the solution” (Driver & Powell, 2016, p. 1). Elementary school students, including ELLs and their EP peers in classrooms taught by CGI-trained teachers, face challenges of varying degrees and from various sources when solving mathematics problems comprised of word-based scenarios. ELLs struggle to match their native EP peers on mathematics achievement assessments comprised of the type of mathematics word problems (Abedi & Lord, 2001; Driver & Powell, 2016; National Center for Education Statistics, 2013) investigated in this study.

Such as the permanence of Base-10 rods and flats as having inherent values of 10 and 100, respectively, while Snap or Base-10 cubes can help with the values in between; or the fact that Snap Cubes can be grouped and color-coded while the Base-10 blocks are typically all the same color.
Concerns have been voiced regarding the inclusion of ELLs in high-stakes, large-scale assessments before they (the students) have developed their academic language proficiency to levels that will allow them to be assessed reliably and validly by assessments as written. This study provides evidence of differential patterns in representational modeling and use of meaning-making tools in a variety of arithmetic word problem types of ELLs. In turn, this study’s findings could be used to help shape future professional development and in-class practice. Solano-Flores et al. (2013) also noted that test developers may use findings from studies investigating semiotic structures of test items to more systematically create test items that are responsive to the validity concerns based on language proficiency expressed by some of the researchers noted earlier. I extend that view to include representational tool use and representational models within that semiotic structures of assessments discussion, so that any updates to future assessments consider also including representational tool use and modeling-based modifications.

Mathematics educators and researchers interested in understanding the challenges faced by students are constantly exploring resources and tools that might increase students’ assessment outcomes without diluting or minimizing the conceptual understanding required to solve these problems. I adopt Fuson, Smith, and Lo Cicero’s (1997) interpretation of assessment as “any teacher-child interaction (directly or by means of a child’s work) that enables the teacher to learn about the child’s thinking to adapt teacher assistance to the state of the child’s conceptual structures” (p. 740), while expanding it to include cognitive
interviews that may serve as extensions of traditional paper-based, individualized assessments.

The available research on ELLs in this area has focused more on their general problem-solving processes and achievement on word problem solving (Ambrose & Molina, 2010; Barwell, 2003, 2005; Bautista, Mulligan, & Mitchelmore, 2009; Cuellar, De La Colina, & Cmajdalka, 2005; Turner & Celedón-Pattichis, 2011). There has been relatively little investigation that has foregrounded representational modeling and the use of representational modeling tools by ELLs during assessment. Though not true of CGI, Gutiérrez and Orellana (2006) also recognize several studies that have investigated how ELLs solve word problems and/or have assumed a deficit model where ELLs are regarded as a monolithic group irrespective of their home language or their mathematics proficiency in this language and/or in English. This study addresses some of the gaps in the research by sharing representations- and assessment-based findings that may be used to inform teachers' instruction of word problem solving for ELLs (Driver & Powell, 2016).

**CGI Problem Types and Solution Strategies**

As noted in Tables 1 and 2 above, there are several well-established mathematics problem types and solution strategies typical in early elementary school grades that CGI has documented and researched over the past several decades. These arithmetic problem types are described, in part, by the
expected actions the student is likely to evidence and/or characteristics of the problem or its prompt: Join (Result, Change, or Start Unknown), Separate (Result, Change, or Start Unknown), Part-Part-Whole (Whole or Part Unknown), Compare (Difference, Compare Quantity, or Referent Unknown), Grouping, and Measurement or Partitive Division (Carpenter, Fennema, Franke, et al., 1999). Similarly, researchers of CGI and other teaching and learning approaches have documented several solution strategies most often observed when students are trying to answer these problem types including: counting-all, counting-on, decomposition, retrieval (Laski, Casey, Yu, Dulaney, Heyman, & Dearing, 2013), direct modeling, counting down, counting down to, and trial and error (Carpenter, Fennema, Franke, et al., 1999). The details of these problem types and solution strategies were specified earlier in this presentation.

There should be a substantial, evident use of representational modeling in the above-referenced strategies. In the counting-on and counting-all strategies, for example, students may first represent addends using one representation system (such as the base-10 or snap cubes), and then count the entirety of the cubes or the rest of the count from the first addend forward, using only their verbal narration of the count. Representational modeling would be seen in a strategy like decomposition, where the student is separating, or decomposing, a given problem into known chunks (such as representing doubles of 3+3 using fingers), before bringing back the decomposed or separated base-10 rod to represent one of the original addends of 13 in the given sum of 13 + 3 (of course, counting-on would also have been a strategy option here).
The arithmetic problem types have several components embedded within their characteristics that support, if not elicit, the representational modeling-based research goals of this study. Join or Separate addition and subtraction problems, for example, call for an explicit combining or separating action embedded in its title, along with various time points (starting time point and quantity, an intermediary time point involving the operation-based change to that starting quantity, that then leads to the final time point where the end quantity is now set) (Carpenter, Fennema, Franke, et al., 1999). These multiple time points afford students ample opportunity to use representation systems to reflect their mathematical thinking on the problem posed. Other problem types such as Part-Part-Whole or Compare problems do not represent an explicit action but still call for an explicit use of representation systems to reflect the implied unknowns of the problem type. Part-part-whole problems have one of three potential components that is initially unknown, but that a familiar use of experienced representational systems might help reveal. Compare problems call for a comparison of one of three different aspects (the original referent set, the compared set, and the difference set) that can also be represented using representational systems and/or tools that may reveal the unknown component of the Compare problem type (Carpenter, Fennema, Franke, et al., 1999). The multiplication and division problem types call for Grouping and Partitive or Measurement Division strategies that, again, implicitly (if not explicitly) direct students to group or partition off quantities that, at least initially, are typically accomplished by direct modeling strategies that then may be followed by other
strategies, such as *counting-all* or *skip counting* (Carpenter, Fennema, Franke, et al., 1999).

CGI is an effective mediator of this archive of problem types and solution strategies because CGI works with teachers to develop an expectation and support protocol for their students when teaching, learning, and using these (and other) strategies to solve these diverse problem types. CGI’s work with younger elementary school students and their mathematics curricula, in turn, makes the use of fingers, base-10 blocks, and snap cubes to be the initial representational modeling tools used by students, as they engage in more direct modeling and then more abstract number facts and counting strategies (Carpenter, Fennema, Franke, et al., 1999). Educators and researchers can analyze the use, sequencing, or misapplication of these tools during individualized assessments to gauge nodes of concept and content comprehension or misunderstanding (using correct responses and corresponding visible solution strategies as a proxy for comprehension, and errors in final responses or solution strategies as evidence of areas in need of further practice and clarification).

**Validity, Reliability, and Bias of Individualized and Whole-Group Assessments**

When any tool that might be potentially useful during an assessment is available to one student but is not available to other students also taking the same assessment, issues of equity and test validity arise. If these threats to an assessment’s validity, reliability, or bias exist, they call into question any findings
from or outcomes based on these assessments. English proficiency is one such (verbal and written) tool, accessible to ELLs and EPs to different degrees. The availability and use of representational systems by all students during an assessment might reduce the bias while increasing the reliability and the validity of standardized mathematics tests. However, if different students have different levels of exposure to and experience with these representational systems and tools, then issues similar to those of English proficiency begin to creep into the results of formative and summative assessments. This study investigated ELLs’ interaction of the use of representational systems and their instructional preparation, while solving word-problem based mathematics assessments through the analysis of videotaped student-interviews recorded during the actual assessment, during a near-real time cognitive interview.

**Cognitive interviews**

Martiniello (2008) recommends the use of cognitive interviews similar to those used in this study to, both, learn how ELLs use and create visual graphic representations while interpreting word problem texts, as well as to help validate ELLs’ test scores in mathematics assessments comprised of word problems. Some of the representations-based data collected through these cognitive interviews may also expose or assuage concerns over validity, reliability, and bias related to this and other mathematics assessments administered in English to ELLs whose English proficiency is, by definition, limited (Abedi, 2004; Durán,
Cognitive interviews made possible the analysis of when and how ELLs evidence and describe their use of representational modeling tools while solving mathematics word problems written in a language in which they may still be developing proficiency.

**Students Enrolled in Mathematics Classrooms Taught by CGI and non-CGI Trained Teachers**

This study focuses on the assessment portion of the “assessing-assisting cycles” (Fuson, Smith, & Lo Cicero, 1997, p. 739), cognizant of the role that teaching plays in a successful assessment experience. Students that participated in this study were enrolled in classes taught by CGI-trained and untrained teachers. CGI professional development (PD) provides an important mechanism through which to share relevant findings regarding representational modeling and tools in their PD activities going forward. Although the students taught by CGI-trained teachers were assessed in individualized settings without their teacher present, it is reasonable to expect that their mathematical problem-solving practices (including those involving representational systems and tools) would draw upon, at least in part, the problem-solving practices that the students experienced in their collaborative classroom settings (Gadge, 2018).
Some teachers and educators believe that there is a sharp distinction between language and mathematics, with mathematics being comprised exclusively of numerical notation, symbolic representations, shapes, and counting, all of which are independent of language and transcend cultural differences (Battey et al., 2013). When mathematical features are regarded as universal, this omits any regard for the language necessary to name and precisely describe those numbers, shapes, and reason through the properties and explanations of solutions (Battey et al., 2013). This assumption of mathematics as universal also restricts mathematics problems to those posed as numeric and symbolic exercises only, devoid of any real-life context or relevance. This perception also helps teachers inadvertently (or purposefully) reinforce ELLs as somehow (cognitively, linguistically, and/or culturally) deficient by focusing on what students cannot do (Moschkovich, 2007b) rather than what they can. When teachers believe that students lack intellectual ability, they may minimize ELLs’ access to language use and quality meaning-creating mathematics opportunities, opting instead to focus on more basic mathematics content, and/or opting to use rote procedural techniques over conceptual techniques, the latter of which might have lead ELLs to discover and develop conceptual mathematics meaning. These choices all but assure that their students will be unable to develop proficiency in the more-complex mathematical thinking that is embodied in the CGI professional development efforts (Battey et al., 2013).
This study positions the use of representational modeling tools by ELLs as effective mediators of tools engaged in object-oriented material activities, and these tools as signs that mediate the construction and understanding of knowledge (Wells, 2007). This study investigates how students’ independent creation and the patterns of representational modeling tools used while solving mathematics word problems by ELLs, communicate a construction and understanding of relevant knowledge called for by the mathematics word problems being solved. The resulting representational models reflect an opportunity to express and manipulate mathematical thinking in a way that words may not (Cobb, Gravemeijer, Yackel, McClain, & Whitenack, 1997; Sfard, 1991; Sfard & Linchevski, 1994), particularly for ELLs. The evident expression and manipulation of mathematical thinking through these multiple representations can then serve as a source of data that may be assessed to gauge student comprehension and guide more effective professional development of pedagogy, particularly by elementary school educators.

**First and Second Graders**

Although elementary school students do not participate in standardized state testing until they reach the third grade, these students still participate in diagnostic, formative, and summative assessments in their classrooms. Their mathematics assessments are comprised of mathematics problem types similar to those found in the assessment used in this study. They are expected to (and
often do) display uses of the varied solution strategies mentioned earlier. Although they are not usually assessed in individualized settings, they are familiar with whole class, small group, and individual assessments (with the collaborative and social supports of the teacher and student’s peers in class).

The assessment used in this study is similar to those typically experienced by the first and second graders that took part in this study. Students’ lifelong mathematical problem-solving practices are being developed at even this young age. It is important to determine if assessment practices that do not foreground representational modeling and tool use should be revised based on studies like this to support more explicit directions for use on assessments. Furthermore, since all participants were given almost exactly the same problems, it is important to keep in mind and account for the difference that a year of formal mathematics learning may afford the second graders over their first grade counterparts. This matters because, while this study evaluates the representational modeling and tool use of all students, it is implicitly doing so with the understanding that this use may, indeed, be moderated by the content knowledge and content and language proficiency the student holds as a result of both more time-on-task in a mathematics learning and assessing setting, and a one-year developmental advantage.
CHAPTER 3
METHODOLOGY

This study investigates the use of representational modeling tools during individualized assessments of selected CGI word problems by ELLs enrolled in 1st and 2nd grade mathematics classrooms taught by teachers who have (and have not) participated in CGI professional development workshops.

This study draws upon a larger study (Schoen, Secada, Dixon, Tazaz, LaVenia, & Childs, 2012-2016) on the effects of Cognitively Guided Instruction on teachers and their students. In this section, I provide an overview of the how the overall student population for the larger study was selected and how this study’s population drawn from the larger; of the assessment instrument that was used in the larger study and of the smaller set of items that are the focus of this study; of student videos on those items were coded; and of the statistical methods that were used on those data.

Study Participants

All the students whose use of representational tool use was analyzed for this study were designated by statutes of the state’s education department as English Language Learners (ELLs). An individualized language proficiency assessment is conducted when any child who, within 20 days of their initial registration into the public school system, answers in the affirmative to any of the
following three questions: (1) Is a language other than English used at home?; (2) Did the student have a first language other than English?; and (3) Does the student most frequently speak a language other than English? (Florida Department of Education, 2011). The Florida Department of Education (2011) directs that, any student who scores within the limited English proficient range as determined by the publisher’s standards on a Department of Education approved aural and oral language proficiency test or scores below the English proficient level on a Department of Education approved assessment in listening and speaking (p.1) is to be classified as an English Language Learner and must be provided with learning support services and resources that respond to this designation.

Not all ELLs are at the same level of English proficiency. Several taxonomies exist that classify these distinctions at discrete levels based on what the ELL can do, including WIDA’s (World-class Instructional Design and Assessment) six levels (“1: Entering”, “2: Emerging”, “3: Developing”, “4: Expanding”, “5: Bridging”, and “6: Reaching”) (Board of Regents of the University of Wisconsin System, 2012, p. 6). Unfortunately, because the demographic data compiled by the school districts for the larger research study team did not include any of the ELLs’ proficiency level details to the research team, the demographic table below does not include these details. The only information provided to the research team of the larger study was an overall ELL vs. non-ELL designation. Furthermore, the district-provided demographic data also did not include any details regarding the language-based instructional programs (i.e., English to Speakers of Other Languages [ESOL], Bilingual Education, etc.) that the ELLs whose videos were
analyzed for this study experienced when in their normal mathematics classrooms.

The 105 public elementary school ELLs selected for this study represented a subset of the 856 cognitive interviews video-recorded by the Replication the CGI experiment in diverse environments team during the second year (academic year of their larger study (Schoen, LaVenia, Champagne, Farina, & Tazaz, 2016). Participants in the assessment, known as the Mathematics Performance and Cognition (MPAC), were selected using a stratified random sampling procedure from an initial sample of 3,681 students whose parent/guardian consented to their participation. These parentally-consented students were enrolled in 22 schools in two diverse school districts in the southeast United States (7 schools from one district and 15 from the other), Although this initial sample was comprised of 1,933 first graders and 1,748 second graders, 440 total first graders (or 22.8% of initial consented group) and 416 total second graders (or 23.8% of the initial consented group) participated in the MPAC student interviews at the end of the 2014-2015 academic year. These students also had completed the Iowa Test of Basic Skills (ITBS; Dunbar et al., 2008) as a pre-test and a post-test at the beginning and at the end of the 2014-2015 academic year; the ITBS was administered in a whole class setting (Schoen, et al., 2016).

The above-referenced cognitive interviews were conducted with students who had completed pretests and whose parental consents for video recording were on file. Two boys and two girls were selected for these interviews. One girl-
boy pair had scored within their below classroom’s median on the pretest; and the second pair scored above their classroom’s (Schoen et al., 2016). Four students could be drawn from almost all of the larger study’s classrooms though some of the classrooms had as few as no students take part in the interview. The interviewers were not informed of either the treatment status of the school (CGI vs. control), nor of the student's pre-test achievement level (below vs above the classroom’s median score).

Of the larger sample's 856 students, most English Language learners (105 of 111 students) were selected for this study. The 6 ELLs’ videos excluded from this analysis were not included because of technical issues with the recording (i.e., video-camera fell during recording; fire alarm during word problem section interrupted recording, etc.). Table 3, an adaptation from Schoen et al. (2016, p.17), presents the detailed demographics funnel data of the initial group of 3,681 participants consented for the larger study, the subset of 856 students eventually video-recorded for the MPAC interview, and the subset of 105 ELLs enrolled in classrooms taught by CGI and non-CGI participants whose representational tool use during the Word Problem portion of their MPAC interview was analyzed for this study. On Table 3, “% of Total” reflects a percentage of the total sample per column (as noted by \( n \) on top row of each main column category). Student characteristic categories are not mutually exclusive. Where data do not sum to the column total, this is most likely due to a binary category (such as ELL vs. non-ELL, or Eligible for free or reduced-price lunch vs. Full-price lunch) whose data provided represents only one of the two possible categories of the binary.
Characteristics noted by “Missing” data represents unreported or unavailable demographic information. While race/ethnicity data were available for the consented and video-recorded samples, those details were not disaggregated for the ELLs selected for analysis for this study.

The final two columns in Table 3 provide demographics data on the Control group in this study. This group of ELLs were enrolled in classrooms where their teachers delivered the school districts’ traditional/conventional forms of instruction. These students joined their Treatment (CGI) group peers in using the state-adopted *GO Math! Florida Student Edition* mathematics textbooks for Grade 1 and Grade 2 (Florida Department of Education, 2012). On the other hand, research by colleagues (Schoen, Secada, & Tazaz, 2015) found that the learning coverage and depth experienced by Treatment and Control students differed in important ways. They found that Treatment students (including ELLs analyzed for this study) were in classrooms whose teachers facilitated the instruction of more advanced content (as defined by the Common Core Mathematics Standards) for all their students and spent more time-on-task than did their Control teacher peers (p. 35). Gadge (2018) also found a difference between the CGI/Treatment and non-CGI/Control students’ opportunities to experience mathematics in noting that the CGI/Treatment teachers were more likely to select and pose word problems (from the same *GO! Math* textbooks) and to listen to all their students’ solutions and thinking (including those of their ELLs) than their non-CGI/Control peers.
Table 3. Demographics of Year 2 consented, interviewed, and ELL students.

<table>
<thead>
<tr>
<th></th>
<th>Total Consented Student Sample (n = 3,681)</th>
<th>Total MPAC Interview Sample (n = 856)</th>
<th>ELLs whose Representational Tool Use in MPAC Interview was analyzed for this study (n = 105)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion of Total n</td>
<td>Proportion of Total n</td>
<td>Proportion of Total n</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.46</td>
<td>1,694</td>
<td>0.50 424</td>
</tr>
<tr>
<td>Female</td>
<td>0.49</td>
<td>1,790</td>
<td>0.50 432</td>
</tr>
<tr>
<td>Missing</td>
<td>0.05</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.53</td>
<td>1,933</td>
<td>0.51 440</td>
</tr>
<tr>
<td>2</td>
<td>0.47</td>
<td>1,748</td>
<td>0.49 416</td>
</tr>
<tr>
<td><strong>Race / Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.04</td>
<td>146</td>
<td>0.06 51</td>
</tr>
<tr>
<td>Black</td>
<td>0.15</td>
<td>562</td>
<td>0.17 144</td>
</tr>
<tr>
<td>White</td>
<td>0.31</td>
<td>1,126</td>
<td>0.33 281</td>
</tr>
<tr>
<td>Other</td>
<td>0.03</td>
<td>93</td>
<td>0.02 21</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.30</td>
<td>1,112</td>
<td>0.32 276</td>
</tr>
<tr>
<td>Missing</td>
<td>0.17</td>
<td>642</td>
<td>0.10 83</td>
</tr>
<tr>
<td><strong>English Language Learners</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>-</td>
<td>-</td>
<td>0.15 16</td>
</tr>
<tr>
<td>Technical Issues</td>
<td>-</td>
<td>-</td>
<td>0.03 3</td>
</tr>
<tr>
<td>Eligible for free or reduced-price lunch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students with disabilities</td>
<td>0.06</td>
<td>231</td>
<td>0.06 52</td>
</tr>
<tr>
<td>Gifted</td>
<td>0.03</td>
<td>97</td>
<td>0.03 28</td>
</tr>
<tr>
<td>Missing</td>
<td>0.17</td>
<td>642</td>
<td>0.10 83</td>
</tr>
</tbody>
</table>

Note: Adapted from Schoen, LaVenia, Champagne, Farina, & Tazaz (2016).
Assessment Instrument

The word problems used for this study was from a broader set of questions which made up three distinct sections: *Counting and Number Screening, Word Problems, and Equations and Expressions*. In a technical report (Schoen et al., 2016), the first-grade version of that instrument had a higher-order math factor composite reliability of 0.91 and a correlation with the ITBS Math Problems (Level 7) subscale of 0.75. The second-grade version of this interview’s higher-order math factor composite reliability was 0.89 and its correlation with the ITBS Math Problems (Level 8) was 0.73. Of the seven grade-specific word problems in this study, two were identical: (i) Separate / Subtraction: Result Unknown; and (ii) Subtraction: Compare Difference Unknown). Most of the remaining Word Problem questions posed to Grade 2 participants differed from those asked of the Grade 1 students. The Grade 2 questions used larger numbers than in Grade 1 in an attempt to increase the difficulty of the problem with the respondent’s age and years of formal schooling. These variations in numerical magnitude were designed to ascertain how the older students made sense of operations involving multi-digit whole numbers (Schoen et al., 2016); a sense-making endeavor that might be reflected in respondents’ use of representational modeling tools.

Although the remaining assessment features (e.g., sequence of sections and some questions, the way questions are posed) were similar enough across grades, the difference in numbers used in most questions and the sequence of
questions for each grade’s assessment precluded any valid and reliable comparison of representational modeling tool uses across participants and groups of different grades.

At the beginning of the Word Problems section, the interviewers reminded students that they could use their fingers and verbal utterances, or any of the five provided representational modeling tools (i.e., drawing paper and markers, snap cubes, and base-10 cubes, rods, and flats) while solving any of the word problems. This reminder was accompanied the interviewer’s by making sure that these representational modeling tools are readily available and accessible to the student participants.

As can be seen in Table 4 below, the Word Problems section was comprised of 7 questions, of standard addition and subtraction or standard multiplication and division (along with an expectation of particular problem-solving strategies that students might engage to solve them), sequenced by increasing difficulty. It is worth noting that the first 5 questions are either identical in both Grade 1 and Grade 2 assessments, or they differ by a numerical magnitude intended to reflect a proportionality between the grade level of the student and the quantities in these problems. Questions 6 and 7, however, are unique to each grade’s assessment.
Table 4. Items (and types and numbers) in the Word Problem section (two questions omitted from analyses in shaded cells below).

<table>
<thead>
<tr>
<th>Gr. __</th>
<th>Q. #</th>
<th>General Question</th>
<th>OPERATION; Problem Type (&amp; Numbers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2</td>
<td>1</td>
<td>There were ### [people doing something after school]. Then # of them [went elsewhere]. How many [people] were still [doing something]?</td>
<td>SEPARATE / SUBTRACTION: Result Unknown (Grades 1 &amp; 2 = 17, 9)</td>
</tr>
<tr>
<td>2 3</td>
<td>1</td>
<td>Someone has # [container of item]. Each [container] has # pieces of [something] in it. How many [pieces of the item does the person] have?</td>
<td>MULTIPLICATION; Grouping (Grade 1 = 6, 5; Grade 2 = 4, 35)</td>
</tr>
<tr>
<td>3 2</td>
<td>1</td>
<td>Someone [performed a task] for # minutes. Someone else worked on her task for ### minutes. How many minutes longer did [the 2nd person] work on her [task] than did [the 1st person]?</td>
<td>SUBTRACTION; Compare Difference Unknown (Grades 1 &amp; 2 = 8, 15)</td>
</tr>
<tr>
<td>4 4</td>
<td>1</td>
<td>Someone has ##(#) [pieces of something]. That person wants to make [something with ## number of pieces on each creation made]. How many creations can the person make if all the pieces are used to make all the creations?</td>
<td>DIVISION; Measurement (Grade 1 = 60, 10, Grade 2 = 110, 10)</td>
</tr>
<tr>
<td>5 5</td>
<td>1</td>
<td>Someone had ## of items somewhere. He received “more” items from somewhere else and put them where the others were, or he wants ## total items. Now, the person has ## items in that place. How many items did the person put onto the place where the items were stored, or does the person need to have the desired total of items?</td>
<td>ADDITION / SUBTRACTION; Join Change Unknown (Missing Addend) (Grade 1 = 15, 24; Grade 2 = 25, 44)</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>Someone had ## [pieces of something]. [Later], he got ## [pieces of the same thing]. How many [pieces] does he have now?</td>
<td>ADDITION; Join Result Unknown (Grade 1 = 49, 56)</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>An item comes in [sets, boxes, groups, etc.] of #. Someone opened # [sets, boxes, groups, etc.] and [did something to them] all. The person needs ## of the items prepared for an event/group/etc. How many more items does the person need to be prepared to have everything he needs?</td>
<td>MULTIPLICATION; Grouping and ADDITION / SUBTRACTION Join Difference Unknown (Grade 2 = 12, 3, 40)</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>An item comes in [sets, boxes, groups, etc.] of #. Someone had # full [sets, boxes, groups, etc.] of the item. The person used # of the items. How many new/unused items does the person have left?</td>
<td>MULTIPLICATION; Grouping and SUBTRACTION Separate Result Unknown (Grade 1 = 12, 3, 4)</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>A place sells a package of # things for # cents. If a person wants to buy ## things, how much will it cost?</td>
<td>DIVISION: Partitive and MULTIPLICATION: Grouping (3, 9, 15)</td>
</tr>
</tbody>
</table>

Note: Adapted from Schoen, LaVenia, Champagne, Farina, & Tazaz (2016).
Cognitive Interview Procedures

At the outset of each cognitive interview, the interviewers informed participating students that they were being video-recorded to help educators and researchers better understand how students solve mathematics problems. The interviewers followed a script of the instructions verbatim provided to the students and of the assessment problems.

Based on years of CGI research and teaching (Carpenter, Fennema, Franke, et al., 1999; Carpenter, Franke, & Levi, 2003), the word problems were asked in order of increasing difficulty, based on a taxonomy of the types of actions or relationships and on the numbers described in the problems (as detailed in Table 1 of Chapter 1). Each response to an assessment problem identified by the student as their final answer served as the interviewer’s cue to proceed to the student’s own explanation portion of the current question (if deemed necessary by the interviewer) before moving on to the next question in the sequence. However, the interviewer was given the option to end the Word Problems section early and move on to the Equations and Expressions section if a student made three consecutive errors in three word problems during the Word Problems section.

This protocol mirrors the real-time item-selection and ability-estimation characteristic of computer adaptive testing (van der Linden & Glas, 2000), and presumes that clear difficulty faced in solving three consecutive word problems
might be a sign of similar unsuccessful attempts with all remaining, even more
difficult, word problems. This was one of several instances where trained
interviewers were encouraged to use their professional judgment to determine
when to move on to the next section or item, whereby minimizing any undue
stress to children who were not progressing well (and may have recognized it).

Regardless of whether a student's response to a problem was correct or
incorrect, interviewers asked follow-up questions that were intended to provide
details on the nature of the student's solution pathway, her/his reasoning, and/or
any unclear explanations or behaviors that the student had given or displayed.
Although no representational-modeling tool use taking place during explanations
were coded for this study, the initial follow-up question for most of the word
problems after the student provided any response was, "How did you get that
answer?" (or a slightly modified version). The follow-up question could also be
specific to a particular representational model or reasoning observed by or
shared with the interviewer: "I saw you using the snap cubes and your fingers to
count. Could you show me exactly how you counted using them?" These follow-
up questions are intended to make more explicit the representations and
reasoning applied by the student, rather than as a way to have students justify or
check their responses. To this end, the interviewers were told to ask additional
questions of the student when their (the students') responses to these initial
follow-up questions did not meaningfully clarify the representational models or
reasoning used, or if the student indicated that they performed most of the
representational modeling and/or reasoning "in (their) head." When students'
representational models and/or reasoning were discernible, the interviewer could state “I see just how you got that answer” and move on to the next question. The purpose of these follow-up questions was to clarify the representational modeling or reasoning that the student used during problem solving; as a result, common pedagogic questions such as “How do you know that your answer is correct?” were not asked. These cognitive interviews provided an opportunity to see and hear how students produced and made use of representational modeling tools and reasoning in their word problem solving.

Each cognitive interview in this study lasted about 45 minutes. As part of a pacing rule established in response to commitments made to parents (and to account for student fatigue resulting from assessments lasting longer than 45 minutes), any student still working on Word Problems 30 minutes after the start of this section’s assessment would be prompted to conclude that word problem and begin the *Equations and Expressions* section of the assessment. The subset of Word Problems analyzed in this study, however, lasted an average of the full 30 minutes.

A webcam was used to record all the interviews. Videos were archived daily to a secure cloud storage server where the larger research team had view access. Although not relevant to the representational modeling data collected from the current study, the interviewers present during the interview wrote notes on paper during the interview that were then shared and archived electronically. These notes were checked for reliability when compared for similarity to notes
made by a second coder who transcribed their own notes on the video to the same shared server.

*Word Problems* were read aloud verbally as the student read the written problem in front of them after being provided, in writing, with one problem per sheet of paper at a time. Students could write on the sheet of paper containing each problem, but these were replaced with the next problem’s sheet of paper upon completion of the prior problem. This detail of the assessment made it possible to adhere to the pacing and section-termination rules described earlier without raising any awareness or alarm in the student because of the interviewer’s decision to skip problems and end a section before all questions were asked.

In all assessments, any student still solving word problems after 30 minutes of working on the entire assessment thus far (including the prior *Number Facts* and *Solving Equations* sections that is not a part of this study) would be asked to begin the subsequent *Equations: True/False* and *Multidigit Computation* sections (also not analyzed for this study because of their decontextualized, mostly-symbolic nature). Similarly, interviewers working with any student still working on the assessment after an hour politely brought the assessment to an end after the completion of the problem at or around the 60th minute. These time limits resulted in incomplete data for some students, as reported in Chapter 4.

The development process of the student interview protocol used in this study is described in further detail elsewhere by the *Replicating the CGI*
Experiment in Diverse Environments team at Florida State University (Schoen et al., 2016).

Coding Representational Modeling Tool Use During Cognitive interviews

The videotaping of each interview made it possible for me to code representational modeling and reasoning in real time. As a result of a pilot study conducted prior to the launch of this more expanded study, the problem-solving phase of the cognitive interviews were the only portions coded and analyzed for this study, as this phase tended to consistently provide evident and ample opportunities for students to engage representational tool uses and for the researcher to code for these; this was less often the case during the problem solving explanation phase of the cognitive interview (if the student was even asked to explain their work or solution by the interviewer).

I started coding each Word Problem when a student made the first audible utterance and/or started to use one of several provided (base-10 cubes, rods, or flat blocks, markers/paper, or snap cube blocks) or existing (i.e., fingers) representational modeling tools. Audible utterances that typically marked the beginning of the Problem Solving included the student reading the problem as written on the assessment sheet or when the interviewer provided the assessment question on the paper and the student began reading it to his/herself. Also, I started coding when the student reached for a marker to start writing or drawing on the provided paper, or starts assembling (or disassembling)
snap cubes, or began manipulating base-10 rods or base-10 unit cubes that might eventually model a quantity specified in the problem. Whether a problem is being read to the student, or the student is reading the problem aloud, or actions in response to the specific details of a given word problem are evident, most students would recognize that an assessment session had begun and would progress towards a solution or a conclusion.

The end of each Assessment (and Problem Question Text) Coverage code is marked by the final audible utterance (most often stating or confirming a final answer to the assessment question) or the purposeful disengagement from the modeling tools. If a student’s final audible utterance or kinesthetic engagement was not evident or purposeful, the interviewer’s acknowledgement of an answer (often with a simple “OK”) also served to mark the end-point of this problem solving phase. At this point, the interviewer typically replaced the current assessment question with a sheet containing the next question. Most students would recognize these actions as the conclusion of one assessment question and the beginning of the next one.

During the second coding pass, the Model/ing Instances evident during the assessment were coded. The coding of these instances segment the assessment exhaustively. Particular care would be taken to account for multiple (double, triple, or more) representational tools used simultaneously, as data of this phenomenon is central to the third research question of interest in this study.

Each problem was coded by its Question number and Grade (so Question
Grade 1, Question 2 Grade 1, etc.... and Question 1 Grade 2 through Question 7 Grade 2). The 7 Representational Tool Use Instances Coded are:

- Base-10 Cube (Value: 1 unit)
- Base-10 Rod (Value: 10 units)
- Base-10 Flat (Value: 100 units)
- Marker Drawing & Paper
- Fingers
- Snap Cubes
- Verbal Response (Self-Talk/Narration)

For the sake of effectiveness, and for potential follow-up studies, each question was also coded to reflect the answer (or lack thereof) finalized by the student:

Correct Solution; Incorrect Solution; No Answer/”I Don’t Know” (Question asked and attempted, but abandoned before a final answer was provided); and Skipped (Question not asked at all).

**Quantitative Approaches to Analyzing Study Data**

The representational modeling evident on video-recordings of students’ solution processes and cognitive interviews while solving the Word Problem questions are accompanied by data on the:

- student’s relative demographics (grade, classroom teacher’s relationship to CGI, English proficiency level);
• tallies of the number of different representational modeling tools used, the aggregate number of distinct uses of these representational modeling tools, and the number of representational tools used simultaneously in pairs;
• though not analyzed for this study: counts of distinct uses of Base-10 blocks, fingers, snap cubes, markers, and verbal response during problem solving;
• As a result, and again, also not analyzed for this study: the sequenced chain of coded representational modeling tools during problem solving.

Why Quantitative/Frequency vs. Qualitative/Descriptive Narratives?

This study seeks to fill a gap in the existing literature in the different kinds of representational tool use by CGI students (as compared to students enrolled in other instructional settings). So the choice, at this juncture, to focus on the distinct frequency of representational tool use exhibited by comparable student groups and/or when solving particular problem types is an effort to establish the basis upon which to pursue future explanatory research that builds on the exploratory ("if it exists") goals of this study.

When these tally data are compared to the students' demographics details, these analyses set the stage for how this study specifically addressed each of the following three focal research questions:

1/VARIETY) How many different representational modeling tools are used…
2/TOTAL) How many total representational modeling tool instances are used…

And
...by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?

To answer these three sets of research questions, 5 x 2 mixed design (within- and between-subject) ANOVAs were conducted for each grade level that examined the effect of the two independent variables at each distinct grade cohort:

(1) the 5 different arithmetic problem types (within-subjects); and

(2) being taught (or not) by a CGI-trained teacher (between-subjects); and

on the following continuous dependent variables:

(1) the total number of different representational modeling tools (out of 7 possible representational tools) used while solving 5 arithmetic questions;

(2) the total, aggregate number of distinct uses of these representational modeling tools; and

(3) the number of simultaneous (2) representational tool uses used at the same time.

Additional analyses were run to determine effect sizes, confidence intervals, and significance of mean differences. Alpha for statistical significance was set of 0.05.

In order to establish coding reliability, 20% of the sample’s coding were coded twice (with at least 3 weeks’ time in between the first and the second
coding runs). The re-coded codes were then compared to assure that there was at least a 90% agreement rate between the pair of codes of the 52 videos whose students’ representational tool use were coded twice.

**Statistical Methodology**

Numerous mixed-design Analyses of Variance (ANOVA) were conducted, since there is at least one within-subject factor (the up to 5 selected Question Types each student could have answered) and one between-subject factor (the Treatment Status designating students as enrolled in classes taught by teachers familiar with the CGI [Treatment] teaching and learning framework, or those teachers not likely to be familiar with CGI, the Control group students). In this mixed design ANOVA, the effect of the representational tool use is inferred through the investigation of the interaction of the question type and the treatment (or control) status. The assumptions for this mixed design ANOVA are similar to those of a between-subject ANOVA (i.e., normality using Q-Q plot, the Kolmogorov-Smirnov Goodness of Fit test, independence of errors, and homogeneity of errors using Levene’s test for Equality of Variances), and within-subject ANOVA (using Mauchly’s Test of Sphericity). For this mixed design ANOVA, the Between-subject factor is comprised of 5 levels for each grade (the 5 questions that ELLs could have answered when assessed that also had sufficient statistical power to be included in these analyses) and the within-
subject factor, comprised of the 2 Treatment Status levels (CGI/Treatment and non-CGI Control).

*IBM SPSS Statistics for Windows* (IBM Corp, Released 2017) was the software used to run the ANOVAs reported in this study. I have since learned about a known issue (Von Hippel, 2004) with SPSS regarding listwise deletion of subjects based on even a single missing data point. This documented shortcoming in the software, resulted in all of the ANOVA results being based on a subsample of the original group of ELLs who responded to each of the 5 questions analyzed for this study. Any student who missed even one of these questions resulted in a listwise deletion by SPSS, so that the total number of students analyzed overall were fewer than the number of those that would have been analyzed at the per-question level if they hadn’t been deleted after missing any one of the five question’s response (and data for analysis). However, since there is an assumption regarding randomization of the original data set, it follows that the data listwise deleted by SPSS was also randomly distributed within the final sample analyzed for this study. While this is not an ideal outcome of the ANOVA analysis, the assumed randomization in both the original data set and the final analyzed data set makes this a tolerable flaw at this juncture. Future, ensuing studies will benefit from using linear mixed-effects models as a better methodology, as these will not completely delete observations from an analysis due to some missing values.

In addition, statistical analyses were conducted on the independent variables (grade level, treatment or control group, and grade-specific question
type) to examine how the observed dependent variable (the number of different representational tools used, of 7 available: Fingers, Markers, Snap Cubes, Base-10 Ones Cubes, Base-10 Tens Rods, and Base-10 Hundreds Flats, and Verbal Responses) changed when any independent variable was different while other independent variables were the same. The mean frequency of different tools used by all the students in the same grade and treatment group while answering each assessment question was calculated.

Given the relatively small sample sizes, the corrected effect size was calculated by using Hedges’ $g$ formula (Hedges’ $g = \frac{M_1 - M_2}{SD_{\text{pooled}}}$, where $M_1-M_2$ is the difference of means, and $SD_{\text{pooled}}$ is the pooled and weighted standard deviation) to measure the effect size, or how much the treatment group statistically differed from the corresponding control group. The pooled standard deviation is calculated as follows: $s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$ where $s_1$ and $s_2$ are the standard deviations of the two samples with sample sizes $n_1$ and $n_2$. We then calculate the Standard Error $se$ of the difference between the two means as: $se(\bar{x}_1 - \bar{x}_2) = s \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$. Cohen (1977) suggested using, with some caveats, the following rule of thumb when interpreting Hedges’ $g$: 0.2 (Small effect; cannot be discerned by the naked eye); 0.5 (Medium effect); 0.8 (Large effect; can be seen with the naked eye).

The Independent Samples 95% Confidence Interval (or accuracy) was calculated in order to determine the range within which we can be 95% confident to find the difference between the two population means. The applicable
Confidence Interval formula for this estimation is: $\mu_1 - \mu_2 = (M_1 - M_2) \pm t s_{(M_1 - M_2)}$, where $M_1$ and $M_2$ are the sample means, $t$ is the $t$-statistic determined by the chosen confidence level (i.e., 95%), and $s_{(M_1 - M_2)}$ is the standard error $= \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}$.

The $p$-value, or the probability of obtaining the observed difference between the samples if the null hypothesis (where the difference would be 0) were true, was also calculated for each question type, student treatment group, and grade. The significance level, or $p$-value, is calculated using the $t$-test, with the value $t$ calculated as: $t = \frac{x_1 - x_2}{se (x_1 - x_2)}$. The $p$-value is the area of the distribution with $n_1 + n_2 - 2$ degrees of freedom, that falls outside $\pm t$. For both the mixed design ANOVAs and the analyses on the independent variables, the $p$-value was set at .05, in accordance with statistical practices in this area.

These statistical analyses were applied to the overall group of students who attempted to answer each of the assessment questions (referred to as "Overall" for each Research Question below) regardless of whether they answered the question correctly, incorrectly, or attempted to answer the question before abandoning it (typically by indicating that they didn’t know or wanted to go on to the next question). As mentioned earlier, the efficacy of varied, total, single, or concurrent representation uses (i.e., engaged when correctly solving a problem). However, in order to provide some efficacy context regarding
participants’ use of representational tools, these statistical analyses were also applied to the subset of respondents who answered the assessment questions correctly (referred to as “Correct Responses” for each Research Question).
CHAPTER 4

RESULTS

Cognitively Guided Instruction (CGI) fosters a teacher’s understanding of their students’ evolving mathematical thinking and how this thinking could serve as the basis for more sophisticated mathematical ideas (Fennema, Carpenter, Franke, Levi, Jacobs, & Empson, 1996). CGI encourages teachers to use their understandings of student reasoning as the basis for creating a new balance between formal or direct instruction on numerical facts and operation algorithms and helping students to develop their own solutions to a variety of mathematics problems (Carpenter et al., 1999). The resulting intuitive processes developed to solve problems involving basic operations often engage representation tool use, modeling, and reasoning worthy of their own investigation.

This chapter presents the results of a study that addresses the following research questions:

1/VARIETY) How many different representational modeling tools are used…

2/TOTAL) How many total representational modeling tool instances are used…

And

3/CONCURRENT) How often are representational modeling tools used individually or simultaneously…
...by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?

Details of the Pilot Study that this study was built on are found in Appendix B.

**Descriptive Statistics**

Table 5 provides details regarding the correct, incorrect, or skipped nature of the various Grade 1 English-learning students’ responses by question type and by treatment or control group. Table 6 summarizes these data for the Grade 2 respondents. Both tables capture a drop-off in the number of total students that attempted to answer each question by half by the end of the assessment. The Correct column in each table represents the number and percentage of total students who attempted to answer the question, and this also shows a decline in percentage of correct respondents from the first through the seventh/last question. The descriptive statistics show that a greater percentage of CGI/Treatment students in Grade 1 (Table 5) answered the majority of the assessment questions correctly than did their non-CGI/Control peers. This observation is far more mixed for Grade 2 (Table 6) CGI/Treatment versus non-CGI/Control correct responses, as each of the two groups seemed to show proficiency in answering some question types that the other group did not.
Table 5.

*Number and percentage of all grade 1 treatment (CGI) and control (non-CGI) respondents for each assessment question type.*

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 1 Question Type / Operation (Numbers)</th>
<th>(T)reatment/CGI or (C)ontrol/non-CGI</th>
<th># of Respondents (% of Total or Correct Resp.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Correct</td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>T</td>
<td>19 (100%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>31 (93.9%)</td>
</tr>
<tr>
<td>2</td>
<td>Multiplication; Grouping (6, 5)</td>
<td>T</td>
<td>19 (100%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>31 (93.9%)</td>
</tr>
<tr>
<td>3</td>
<td>Subtraction; Compare Difference Unknown (8, 15)</td>
<td>T</td>
<td>19 (100%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>32 (97.0%)</td>
</tr>
<tr>
<td>4</td>
<td>Division; Measurement (60, 10)</td>
<td>T</td>
<td>15 (78.9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>24 (72.7%)</td>
</tr>
<tr>
<td>5</td>
<td>Change Unknown (Missing Addend) (15, 24)</td>
<td>T</td>
<td>11 (57.9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>12 (36.4%)</td>
</tr>
</tbody>
</table>

Table 6.

*Number and percentage of all grade 2 treatment (CGI) and control (non-CGI) respondents for each assessment question type.*

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 2 Question Type / Operation (Numbers)</th>
<th>(T)reatment/CGI or (C)ontrol/non-CGI</th>
<th># of Respondents (% of Total or Correct Resp.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Correct</td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>T</td>
<td>25 (100.0%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>28 (100.0%)</td>
</tr>
<tr>
<td>2</td>
<td>Subtraction; Compare Difference Unknown (8, 15)</td>
<td>T</td>
<td>25 (100.0%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>28 (100.0%)</td>
</tr>
<tr>
<td>3</td>
<td>Multiplication; Grouping (4, 35)</td>
<td>T</td>
<td>25 (100.0%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>28 (100.0%)</td>
</tr>
<tr>
<td>4</td>
<td>Division; Measurement (110, 10)</td>
<td>T</td>
<td>24 (96.0%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>26 (92.9%)</td>
</tr>
<tr>
<td>5</td>
<td>Change Unknown (Missing Addend) (25, 44)</td>
<td>T</td>
<td>13 (52.0%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>23 (82.1%)</td>
</tr>
</tbody>
</table>

A review of the questions asked on each grade’s assessment (as specified in Table 4, Chapter 3), coupled with a review of the total and correct
response rates by the ELLs (as noted in Tables 5 and 6 above), provide support for my decision to omit Questions 6 and 7 from any inferential analyses due to a lack of data (and therefore statistical power), and the limitation of the grade-specific nature of these final two (presumably most difficult) questions of each grade’s assessment. Tables 7, 8, 9, and 10 below detail the mean, sample size, and standard deviation of the representational tool use by Grade 1 and Grade 2 Treatment and Control ELLs, by research questions and response type (all questions, questions answered correctly, and those answered incorrectly).

**Inferential Statistics**

This section details the result of inferential statistics including (a) mixed design Analyses of Variance (ANOVAs), to protect against spurious results; as well as (b) Hedges’ $g$ effect sizes, 95% Confidence Intervals, and $p$-values of mean differences of the relevant independent variables.

**Research Question 1 - Overall**

Research question 1 asked, *How many different representational modeling tools are used by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>19</td>
<td>3.16 1.26</td>
<td>5.33 3.03</td>
<td>4.21 3.24</td>
<td>2.83 1.34</td>
</tr>
<tr>
<td>1 RIGHT</td>
<td>14</td>
<td>3.00 1.11</td>
<td>5.46 3.20</td>
<td>4.00 2.22</td>
<td>3.38 1.30</td>
</tr>
<tr>
<td>WRONG</td>
<td>5</td>
<td>3.60 1.67</td>
<td>5.00 2.83</td>
<td>4.80 5.54</td>
<td>1.75 0.50</td>
</tr>
<tr>
<td>ALL</td>
<td>19</td>
<td>2.84 1.38</td>
<td>6.50 5.23</td>
<td>4.21 4.29</td>
<td>3.69 2.81</td>
</tr>
<tr>
<td>2 RIGHT</td>
<td>7</td>
<td>2.86 1.07</td>
<td>8.86 4.71</td>
<td>4.57 3.36</td>
<td>4.17 3.54</td>
</tr>
<tr>
<td>WRONG</td>
<td>12</td>
<td>2.83 1.59</td>
<td>5.00 4.94</td>
<td>4.00 4.88</td>
<td>3.29 2.21</td>
</tr>
<tr>
<td>ALL</td>
<td>19</td>
<td>2.95 1.39</td>
<td>4.42 2.69</td>
<td>3.58 2.52</td>
<td>2.00 0.94</td>
</tr>
<tr>
<td>3 RIGHT</td>
<td>6</td>
<td>2.83 1.47</td>
<td>4.67 2.58</td>
<td>3.33 1.21</td>
<td>2.00 1.41</td>
</tr>
<tr>
<td>WRONG</td>
<td>13</td>
<td>3.00 1.41</td>
<td>4.31 2.84</td>
<td>3.69 2.98</td>
<td>2.00 0.93</td>
</tr>
<tr>
<td>ALL</td>
<td>15</td>
<td>2.40 1.18</td>
<td>4.53 4.78</td>
<td>3.50 3.48</td>
<td>2.50 1.87</td>
</tr>
<tr>
<td>4 RIGHT</td>
<td>5</td>
<td>1.60 0.55</td>
<td>2.40 2.19</td>
<td>2.25 2.50</td>
<td>1.50 0.71</td>
</tr>
<tr>
<td>WRONG</td>
<td>9</td>
<td>2.56 1.01</td>
<td>5.33 5.70</td>
<td>3.67 3.87</td>
<td>3.00 2.16</td>
</tr>
<tr>
<td>IDK</td>
<td>1</td>
<td>5.00 %</td>
<td>8.00 %</td>
<td>7.00 %</td>
<td>- %</td>
</tr>
<tr>
<td>ALL</td>
<td>12</td>
<td>2.45 1.04</td>
<td>4.45 2.70</td>
<td>3.82 2.48</td>
<td>1.00 0.00</td>
</tr>
<tr>
<td>5 RIGHT</td>
<td>8</td>
<td>3.00 0.58</td>
<td>5.43 2.57</td>
<td>4.57 2.64</td>
<td>1.00 0.00</td>
</tr>
<tr>
<td>WRONG</td>
<td>4</td>
<td>1.50 1.00</td>
<td>2.75 2.22</td>
<td>2.50 1.73</td>
<td>1.00 %</td>
</tr>
</tbody>
</table>

Note. % = “DIV/0!” Error when < 2 numeric values used when calculating s.d. “IDK” = “I Don’t Know” (question abandoned after trying to solve it; does not include questions skipped/not asked).
Table 8.

*Grade 1 Control (non-CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type.*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>31</td>
<td>2.97</td>
<td>1.22</td>
<td>5.71</td>
<td>4.95</td>
</tr>
<tr>
<td>1 RIGHT</td>
<td>19</td>
<td>3.05</td>
<td>1.39</td>
<td>5.74</td>
<td>3.84</td>
</tr>
<tr>
<td>WRONG</td>
<td>12</td>
<td>2.83</td>
<td>0.94</td>
<td>5.67</td>
<td>4.92</td>
</tr>
<tr>
<td>ALL</td>
<td>31</td>
<td>2.61</td>
<td>1.15</td>
<td>5.45</td>
<td>4.23</td>
</tr>
<tr>
<td>2 RIGHT</td>
<td>4</td>
<td>3.00</td>
<td>0.82</td>
<td>6.25</td>
<td>2.67</td>
</tr>
<tr>
<td>WRONG</td>
<td>25</td>
<td>2.36</td>
<td>0.99</td>
<td>3.92</td>
<td>3.08</td>
</tr>
<tr>
<td>IDK</td>
<td>2</td>
<td>5.00</td>
<td>0.00</td>
<td>23.00</td>
<td>21.00</td>
</tr>
<tr>
<td>ALL</td>
<td>32</td>
<td>2.34</td>
<td>1.07</td>
<td>3.97</td>
<td>3.10</td>
</tr>
<tr>
<td>3 RIGHT</td>
<td>7</td>
<td>2.43</td>
<td>0.79</td>
<td>4.86</td>
<td>4.50</td>
</tr>
<tr>
<td>WRONG</td>
<td>25</td>
<td>2.32</td>
<td>1.14</td>
<td>3.72</td>
<td>2.76</td>
</tr>
<tr>
<td>IDK</td>
<td>2</td>
<td>5.00</td>
<td>0.00</td>
<td>23.00</td>
<td>21.00</td>
</tr>
<tr>
<td>ALL</td>
<td>24</td>
<td>2.33</td>
<td>1.17</td>
<td>4.33</td>
<td>3.21</td>
</tr>
<tr>
<td>4 RIGHT</td>
<td>4</td>
<td>2.25</td>
<td>0.96</td>
<td>2.75</td>
<td>2.25</td>
</tr>
<tr>
<td>WRONG</td>
<td>17</td>
<td>2.29</td>
<td>1.26</td>
<td>4.59</td>
<td>3.41</td>
</tr>
<tr>
<td>IDK</td>
<td>3</td>
<td>2.67</td>
<td>1.15</td>
<td>5.00</td>
<td>3.33</td>
</tr>
<tr>
<td>ALL</td>
<td>12</td>
<td>2.45</td>
<td>1.04</td>
<td>4.45</td>
<td>3.82</td>
</tr>
<tr>
<td>5 RIGHT</td>
<td>8</td>
<td>3.00</td>
<td>0.58</td>
<td>5.43</td>
<td>4.57</td>
</tr>
<tr>
<td>WRONG</td>
<td>4</td>
<td>1.50</td>
<td>1.00</td>
<td>2.75</td>
<td>2.50</td>
</tr>
</tbody>
</table>

*Note.* % = “DIV/0!” Error when < 2 numeric values used when calculating s.d. “IDK” = “I Don’t Know” (question abandoned after trying to solve it; does not include questions skipped/not asked).
Table 9.
Grade 2 Treatment (CGI) ELLs’ mean n and s.d. of representational tool use by research question, question number & response type.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>25</td>
<td>3.00 1.32</td>
<td>5.40 3.23</td>
<td>4.16 2.67</td>
<td>2.08 1.04</td>
</tr>
<tr>
<td>1 RIGHT</td>
<td>21</td>
<td>2.95 1.43</td>
<td>5.43 3.46</td>
<td>4.19 2.80</td>
<td>2.20 1.14</td>
</tr>
<tr>
<td>WRONG</td>
<td>4</td>
<td>3.25 0.50</td>
<td>5.25 1.89</td>
<td>4.00 2.16</td>
<td>1.67 0.58</td>
</tr>
<tr>
<td>ALL</td>
<td>25</td>
<td>2.72 1.02</td>
<td>5.24 3.61</td>
<td>3.96 2.47</td>
<td>2.17 1.59</td>
</tr>
<tr>
<td>2 RIGHT</td>
<td>11</td>
<td>2.73 1.19</td>
<td>5.73 4.05</td>
<td>4.18 2.72</td>
<td>2.80 2.17</td>
</tr>
<tr>
<td>WRONG</td>
<td>14</td>
<td>2.71 0.91</td>
<td>4.86 3.32</td>
<td>3.79 2.33</td>
<td>1.71 0.95</td>
</tr>
<tr>
<td>ALL</td>
<td>25</td>
<td>3.00 1.22</td>
<td>7.56 4.57</td>
<td>6.64 4.35</td>
<td>1.91 1.38</td>
</tr>
<tr>
<td>3 RIGHT</td>
<td>9</td>
<td>3.11 1.45</td>
<td>9.00 5.22</td>
<td>8.33 5.05</td>
<td>1.20 0.45</td>
</tr>
<tr>
<td>WRONG</td>
<td>15</td>
<td>2.93 1.16</td>
<td>6.93 4.20</td>
<td>5.80 3.84</td>
<td>2.50 1.64</td>
</tr>
<tr>
<td>IDK</td>
<td>1</td>
<td>3.00 %</td>
<td>4.00 %</td>
<td>4.00 %</td>
<td>-</td>
</tr>
<tr>
<td>ALL</td>
<td>24</td>
<td>2.88 1.30</td>
<td>6.92 5.63</td>
<td>5.75 4.35</td>
<td>1.92 1.24</td>
</tr>
<tr>
<td>4 RIGHT</td>
<td>8</td>
<td>3.13 0.99</td>
<td>7.25 6.04</td>
<td>6.00 4.69</td>
<td>2.00 1.15</td>
</tr>
<tr>
<td>WRONG</td>
<td>15</td>
<td>2.73 1.49</td>
<td>6.40 5.62</td>
<td>5.27 4.22</td>
<td>2.00 1.41</td>
</tr>
<tr>
<td>IDK</td>
<td>1</td>
<td>3.00 %</td>
<td>12.00 %</td>
<td>11.00 %</td>
<td>1.00 %</td>
</tr>
<tr>
<td>ALL</td>
<td>13</td>
<td>3.00 1.08</td>
<td>6.92 6.18</td>
<td>6.08 5.62</td>
<td>1.43 0.79</td>
</tr>
<tr>
<td>5 RIGHT</td>
<td>10</td>
<td>3.00 0.94</td>
<td>5.60 2.41</td>
<td>4.90 2.60</td>
<td>1.40 0.89</td>
</tr>
<tr>
<td>WRONG</td>
<td>3</td>
<td>3.00 1.73</td>
<td>11.33 12.86</td>
<td>10.00 11.36</td>
<td>1.50 0.71</td>
</tr>
</tbody>
</table>

*Note. % = “DIV/0!” Error when < 2 numeric values used when calculating s.d. “IDK” = “I Don’t Know” (question abandoned after trying to solve it; does not include questions skipped/not asked).*
Table 10.

Grade 2 Control (non-CGI) ELLs' mean n and s.d. of representational tool use by research question, question number & response type.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>s.d.</td>
<td>mean</td>
<td>s.d.</td>
<td>mean</td>
</tr>
<tr>
<td>ALL</td>
<td>28</td>
<td>2.79</td>
<td>1.17</td>
<td>4.86</td>
<td>3.15</td>
</tr>
<tr>
<td>1 RIGHT</td>
<td>25</td>
<td>2.84</td>
<td>1.18</td>
<td>5.04</td>
<td>3.21</td>
</tr>
<tr>
<td>WRONG</td>
<td>3</td>
<td>2.33</td>
<td>1.15</td>
<td>3.33</td>
<td>2.52</td>
</tr>
<tr>
<td>ALL</td>
<td>28</td>
<td>2.73</td>
<td>1.12</td>
<td>5.04</td>
<td>4.05</td>
</tr>
<tr>
<td>2 RIGHT</td>
<td>17</td>
<td>2.47</td>
<td>1.01</td>
<td>4.59</td>
<td>3.86</td>
</tr>
<tr>
<td>WRONG</td>
<td>9</td>
<td>3.22</td>
<td>1.20</td>
<td>7.00</td>
<td>4.15</td>
</tr>
<tr>
<td>IDK</td>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ALL</td>
<td>28</td>
<td>2.89</td>
<td>1.34</td>
<td>6.57</td>
<td>6.06</td>
</tr>
<tr>
<td>3 RIGHT</td>
<td>10</td>
<td>2.90</td>
<td>1.37</td>
<td>7.80</td>
<td>6.91</td>
</tr>
<tr>
<td>WRONG</td>
<td>17</td>
<td>3.00</td>
<td>1.32</td>
<td>6.18</td>
<td>5.66</td>
</tr>
<tr>
<td>IDK</td>
<td>1</td>
<td>1.00</td>
<td>%</td>
<td>1.00</td>
<td>%</td>
</tr>
<tr>
<td>ALL</td>
<td>26</td>
<td>2.52</td>
<td>1.00</td>
<td>4.38</td>
<td>3.06</td>
</tr>
<tr>
<td>4 RIGHT</td>
<td>8</td>
<td>2.00</td>
<td>1.07</td>
<td>2.75</td>
<td>1.28</td>
</tr>
<tr>
<td>WRONG</td>
<td>16</td>
<td>2.75</td>
<td>0.93</td>
<td>5.56</td>
<td>3.35</td>
</tr>
<tr>
<td>IDK</td>
<td>2</td>
<td>3.00</td>
<td>%</td>
<td>1.50</td>
<td>%</td>
</tr>
<tr>
<td>ALL</td>
<td>23</td>
<td>2.96</td>
<td>1.07</td>
<td>6.59</td>
<td>4.13</td>
</tr>
<tr>
<td>5 RIGHT</td>
<td>9</td>
<td>2.67</td>
<td>1.22</td>
<td>6.89</td>
<td>6.13</td>
</tr>
<tr>
<td>WRONG</td>
<td>14</td>
<td>3.14</td>
<td>0.95</td>
<td>6.07</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Note. % = “DIV/0!” Error when < 2 numeric values used when calculating s.d. “IDK” = “I Don’t Know” (question abandoned after trying to solve it; does not include questions skipped/not asked).
ANOVA: Grade 1 Variety of Representational Tools Used

An insignificant Mauchly’s W of .64 indicated that a within-subject factor, Question Type (from a Variety of Representational Tools Used perspective), met the sphericity assumption, $\chi^2(9) = 7.40, p = .60$. Also, the insignificant Levene’s tests showed that all repeated measures had the same error variances, $F(1,18) = .61, p = .45$ for Question #1 (Type: “Separate / Subtraction: Result Unknown”); $F(1,18) < .001, p = .99$ for Question #2 (Type: “Multiplication: Grouping”); $F(1,18) = 1.04, p = .32$ for Question #3 (Type: “Subtraction; Compare Difference Unknown”); $F(1,18) = .88, p = .36$ for Question #4 (Type: “Division: Measurement”); and $F(1,18) = .19, p = .67$ for Question #5 (Type: “Addition/Subtraction: Join Change Unknown (Missing Addend)”.

As shown in Table 11 below, a 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Variety of Representational Tools Used by Grade 1 students showed that the main effects of Question Type is statistically significant, $F(4,72) = 3.39, p < .05, \eta^2_p = .16$ is statistically significant, while Treatment Status is not statistically significant, $F(1,18) = .32, p = 0.58, \eta^2_p = .02$. Their two-way interaction, however, was found to be statistically significant, $F(4,72) = 2.81, p < .05, \eta^2_p = .14$. 
Table 11.

*Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Variety of Representational Tools Used by Grade 1 students.*

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>(\eta^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Type</td>
<td>11.57</td>
<td>4</td>
<td>2.89</td>
<td>3.39</td>
<td>&lt;.05*</td>
<td>.16</td>
</tr>
<tr>
<td>Question Type * Treatment Status</td>
<td>9.61</td>
<td>4</td>
<td>2.40</td>
<td>2.81</td>
<td>&lt;.05*</td>
<td>.14</td>
</tr>
<tr>
<td>Error (Question Type)</td>
<td>61.50</td>
<td>72</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Status</td>
<td>1.11</td>
<td>1</td>
<td>1.11</td>
<td>.32</td>
<td>.58</td>
<td>.02</td>
</tr>
<tr>
<td>Error</td>
<td>63.64</td>
<td>18</td>
<td>3.54</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * \(p < 0.05\).

Since the interaction of Question Type and Treatment Status was shown to be significant, a pairwise comparison with Bonferroni adjustment was run on the between-subjects factor, Treatment Status. This pairwise comparison with Bonferroni adjustment showed that the mean difference of variety of representational tools used by Grade 1 students was significant (\(M_{diff} = 1.26, p < 0.05\)) when CGI/Treatment students answered Question #5 (Type: “Addition/Subtraction: Join Change Unknown (Missing Addend)”) versus when their non-CGI/Control peers answered the same question, with a higher average number of different representational tools used by CGI/Treatment students when answering Question #5 than by their non-CGI/Control peers answering the same question (3.82 vs. 2.56, respectively).
Effect Size: Grade 1 Variety of Representational Tools Used

Table 12 summarizes the results of these calculations on the variety of representational tools used (of 7 available representational tools). While there was no statistically significant difference between the variety of representational tools used by Grade 1 Treatment and Control group students for most question types, there was a statistical difference (Hedges’ $g = 1.152$, 95% Confidence Interval for Differences = 0.336 --2.391, $p < .05$) in the variety of representational tools used by Grade 1 Treatment and Control students when solving “Addition/Subtraction: Join Change Unknown (Missing Addend)” (Question #5) problem types (with a pair of 2-digit numbers less than 50).

Table 12.

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 1 Question Type / Operation (Numbers)</th>
<th>Average # of Different Representational Tools Used (of 7 possible tools)</th>
<th>Hedges’ $g$</th>
<th>Independent $p$-value for mean difference (2-tailed $t$-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>Treatment (CGI) 3.16</td>
<td>Control (non-CGI) 2.97</td>
<td>0.154</td>
</tr>
<tr>
<td>2</td>
<td>Multiplication: Grouping (6, 5)</td>
<td>2.84</td>
<td>2.61</td>
<td>0.185</td>
</tr>
<tr>
<td>3</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.95</td>
<td>2.34</td>
<td>0.503</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (60,10)</td>
<td>2.40</td>
<td>2.33</td>
<td>0.057</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (15, 24)</td>
<td>3.82</td>
<td>2.45</td>
<td>1.152</td>
</tr>
</tbody>
</table>

Note: * $p < 0.05$.
ANOVA: Grade 2 Variety of Representational Tools Used

As shown in Table 13, the mixed-design ANOVA executed to try to provide statistical insight on Research Question #1 (regarding Treatment and Control group ELLs’ use of a variety of representational tools) did not result in any statistically significant findings.

Table 13.
Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Variety of Representational Tools Used by Grade 2 students.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Type</td>
<td>3.28</td>
<td>4</td>
<td>.82</td>
<td>1.51</td>
<td>.20</td>
<td>.04</td>
</tr>
<tr>
<td>Question Type * Treatment Status</td>
<td>.83</td>
<td>4</td>
<td>.21</td>
<td>.39</td>
<td>.82</td>
<td>.01</td>
</tr>
<tr>
<td>Error (Question Type)</td>
<td>71.54</td>
<td>132</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Status</td>
<td>1.53</td>
<td>1</td>
<td>1.53</td>
<td>.37</td>
<td>.55</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td>136.78</td>
<td>33</td>
<td>4.15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Effect Size: Grade 2 Variety of Representational Tools Used

Table 14 summarizes the results of these calculations on the variety of representational tools used, or the relationship between the average number of different representational tools used (of 7 available representational tools) and the various independent demographic variables for Grade 2 students. Unlike the
Grade 1 findings in this criteria, no statistically significant difference was found between the variety of representational tools used by Grade 2 Treatment and Grade 2 Control group students for any of the question types asked.

Table 14.

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 2 Question Type / Operation (Numbers)</th>
<th>Average # of Different Representational Tools Used (of 7 possible tools)</th>
<th>Hedges' g</th>
<th>Independent p-value for mean difference (2-tailed t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>3.00</td>
<td>2.79</td>
<td>0.169</td>
</tr>
<tr>
<td>2</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.72</td>
<td>2.73</td>
<td>0.009</td>
</tr>
<tr>
<td>3</td>
<td>Multiplication: Grouping (4, 35)</td>
<td>3.00</td>
<td>2.89</td>
<td>0.085</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (110, 10)</td>
<td>2.88</td>
<td>2.52</td>
<td>0.312</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (25, 44)</td>
<td>3.00</td>
<td>2.96</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Research Question 2 - Overall

Research question 2 asked, How many total representational modeling tool instances are used by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?
For this research question (and in accordance with how data were recorded), “instances” of representational tools used are noted whenever a student uses any of the 7 provided/available representational tools (i.e., Base-10 blocks, cubes, or rods; snap cubes; markers and paper; verbal expression [as self-explanation, narration, etc.]; and fingers) observed to be used individually, or in conjunction in pairs, triples, or even tetrads. To clarify: when a student is noted to have exhibited 5 total representational tool instances, this representational total use may be a combination of 5 distinct uses of one representation tool used at a time and consecutively (with the possibility of repeating one or more tools used distinctly), or it could be a combination of, for example, 2 unique, distinct representational tools used, 2 additional distinct instances of concurrent pairs of tools used (e.g., fingers and Base-10 cubes, and then markers and paper and an accompanying verbal expression), and perhaps even a single observation of three representational tools used simultaneously (such as Base-10 rods, fingers, and an accompanying, related verbal expression). Both of these scenarios would be tallied as 5 total representational tool instances used, even though they look different from each other and engage completely different sets of tools. In short, an instance should be regarded as distinct enacted expressions of representational tool use.
ANOVA: Grade 1 Total # of Representational Tools Used

An insignificant Mauchly’s W of .40 indicated that a within-subject factor, Question Type (from a Total # of Representational Tools Used perspective), met the sphericity assumption ($\chi^2(9) = 14.92, p = .10$). Also, the insignificant Levene’s tests showed that all repeated measures had the same error variances ($F(1,18) = .05, p = .83$ for Question #1 (Type: “Separate / Subtraction: Result Unknown”); $F(1,18) = 2.94, p = .10$ for Question #2 (Type: “Multiplication: Grouping”); $F(1,18) = .28, p = .60$ for Question #3 (Type: “Subtraction: Compare Difference Unknown”); $F(1,18) = .27, p = .61$ for Question #4 (Type: “Division: Measurement”); $F(1,18) = 2.27, p = .15$ for Question #5 (Type: “Addition/Subtraction: Join Change Unknown (Missing Addend)”)). As shown in Table 15 below, A 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 1 students showed that the main effects of Question Type, $F(4,72) = 3.39, p < .05, \eta^2_p = 0.16$, is significant, while Treatment Status, $F(1,18) = .31, p = 0.58, \eta^2_p = 0.02$, is not statistically significant. Also, their two-way interaction was found to not be statistically significant, $F(4,72) = 1.32, p = .27, \eta^2_p = .07$. 

Table 15.
Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 1 students.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Type</td>
<td>164.67</td>
<td>4</td>
<td>41.17</td>
<td>3.39</td>
<td>&lt;.05*</td>
<td>0.16</td>
</tr>
<tr>
<td>Question Type * Treatment Status</td>
<td>64.34</td>
<td>4</td>
<td>16.09</td>
<td>1.32</td>
<td>.27</td>
<td>.07</td>
</tr>
<tr>
<td>Error (Question Type)</td>
<td>874.60</td>
<td>72</td>
<td>12.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Status</td>
<td>15.13</td>
<td>1</td>
<td>15.13</td>
<td>.31</td>
<td>.58</td>
<td>.02</td>
</tr>
<tr>
<td>Error</td>
<td>875.38</td>
<td>18</td>
<td>48.63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * $p < 0.05$.

Since the two-way interaction in this ANOVA was found to be not significant, no additional effect size analysis for the interactions was necessary.

ANOVA: Grade 2 Total # of Representational Tools Used

A significant Mauchly’s $W$ of .55 indicated that a within-subject factor, Question Type (from a Total # of Representational Tools Used perspective), did not meet the sphericity assumption ($\chi^2(9) = 18.76, p < .05$), calling on the use of the Huynh-Feldt Epsilon ($\varepsilon$) adjustment for degrees of freedom in this mixed-design ANOVA. Also, the insignificant Levene’s tests showed that all repeated measures had the same error variances ($F(1,33) = .26, p = .61$ for Question #1 (Type: “Separate / Subtraction: Result Unknown”); $F(1,33) = .36, p = .55$ for Question #2 (Type: “Subtraction; Compare Difference Unknown”); $F(1,33) = .34,
$p = .57$ for Question #3 (Type: “Multiplication: Grouping”); $F(1,33) = .76, p = .39$ for Question #4 (Type: “Division: Measurement”); and $F(1,33) = .20, p = .66$ for Question #5 (Type: “Addition/Subtraction: Join Change Unknown (Missing Addend)”). As shown in Table 17 below, A 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 2 students showed that the main effects of Question Type, $F(3.63,119.93) = 3.47, p < .05, \eta^2 = 0.10$, is significant, while Treatment Status, $F(1,33) = 1.03, p = 0.32, \eta^2 = 0.03$, is not statistically significant. Also, their two-way interaction was found to not be statistically significant, $F(3.63,119.93) = 0.33, p = .84, \eta^2 = .01$.

Table 16.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question Type</td>
<td>210.18</td>
<td>3.63</td>
<td>57.83</td>
<td>3.47</td>
<td>&lt; .05*</td>
<td>.10</td>
</tr>
<tr>
<td>Question Type * Treatment</td>
<td>20.06</td>
<td>3.63</td>
<td>5.52</td>
<td>.33</td>
<td>.84</td>
<td>.01</td>
</tr>
<tr>
<td>Error (Question Type)</td>
<td>2000.99</td>
<td>119.93</td>
<td>16.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Status</td>
<td>40.91</td>
<td>1</td>
<td>40.91</td>
<td>1.03</td>
<td>.32</td>
<td>.03</td>
</tr>
<tr>
<td>Error</td>
<td>1308.83</td>
<td>33</td>
<td>39.66</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * $p < 0.05$. 

Results from 5 (Question Type) by 2 (Treatment Status) mixed-design ANOVA of the Total # of Representational Tools Used by Grade 2 students.
A pairwise comparison with Bonferroni adjustment run showed that the mean difference of Total # of representational tools used by Grade 2 students was significant (\(M_{\text{diff}} = 3.01, p < 0.05\), for both) when they answered Question #3 (Type: “Multiplication: Grouping”) versus when they answered both Question #1 (Type: “Separate / Subtraction: Result Unknown”) and Question #2 (Type: “Subtraction; Compare Difference Unknown”), with a higher average number of total representational tools used by students when answering Question #3 than Question #1, and Question #3 than Question #2 (8.14 vs. 5.13 for both pairwise comparisons). No other post-hoc of within-subjects or between-subjects analyses provided any significant findings.

**Research Question 3 – Overall**

Research question 3 asked, *How often are representational modeling tools used individually or simultaneously by 1st and 2nd Grade ELLs enrolled in classes taught by teachers trained in (i.e., Treatment) or not trained in (i.e., Control) Cognitively Guided Instruction (CGI) pedagogy when solving various arithmetic word problem types?*

To address this research question, statistical calculations similar to those conducted for the other research questions were carried out to identify any statistically significant relationships between the total number of representational tools used individually or simultaneously (two at a time), and the various independent demographic variables considered thus far. Both Grade 1 and
Grade 2 Treatment and Control group students were observed to use simultaneous (two) representational tools differently from how they used the individual representational tools.

**ANOVA: Grade 1 # of Double/Concurrent Representational Tools Used**

The mixed-design ANOVA executed to try to provide statistical insight on Research Question #3 (regarding Treatment and Control group ELLs’ use of double/concurrent representational tools) did not result in any such insight due to the excessive number of missing data.

**Effect Size: Grade 1 # of Double/Concurrent Representational Tools Used**

Table 17 summarizes the data calculated for Research Question 3 regarding Grade 1 students’ simultaneous use of representational tools when answering various question types. Three of the five question types asked (and analyzed) of Grade 1 students evidenced a statistically significant difference between Grade 1 Treatment group students’ use of simultaneous (two) representational tools and Grade 1 Control group students’ use of simultaneous (two) representational tools: “Addition/Subtraction: Join Change Unknown (Missing Addend)” problem type (Hedges’ $g = 1.711$, 95% Confidence Interval for Differences $= 1.022 – 2.978$, $p < .01$); “Separate / Subtraction: Result Unknown” problem type (Hedges’ $g = 0.713$, 95% Confidence Interval for Differences $=$
0.174 – 1.779, *p < .05*; and “Multiplication: Grouping” problem type (Hedges’ *g* = 0.613, 95% Confidence Interval for Differences = 0.065 – 2.858, *p < .05*). In each of these instances, the findings showing the CGI students using a greater average number of multiple, simultaneous representational tools than did their non-CGI peers on 3 of the 5 questions analyzed, suggesting that CGI students are using more complex representational modeling than their non-CGI peers.

Table 17.

*Total frequency of use of concurrent (two) representational tools by grade 1 treatment (CGI) and control (non-CGI) respondents for each assessment question type.*

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 1 Question Type / Operation (Numbers)</th>
<th>Average # of Concurrent (Two) Representational Tools Used</th>
<th>Hedges’ <em>g</em></th>
<th>Independent Samples 95% Confidence Interval for Mean Difference (2-tailed <em>t</em>-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>2.83 1.86</td>
<td>0.713</td>
<td>0.174 – 1.779 <em>p &lt; .05</em></td>
</tr>
<tr>
<td>2</td>
<td>Multiplication: Grouping (6, 5)</td>
<td>3.69 2.33</td>
<td>0.613</td>
<td>0.065 – 2.858 <em>p &lt; .05</em></td>
</tr>
<tr>
<td>3</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.00 1.82</td>
<td>0.243</td>
<td>-0.257 – 0.620 <em>p &lt; .05</em></td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (60, 10)</td>
<td>2.50 1.88</td>
<td>0.310</td>
<td>-0.720 – 1.970 <em>p &lt; .05</em></td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (15, 24)</td>
<td>3.00 1.00</td>
<td>1.711</td>
<td>1.022 – 2.978 <strong>p &lt; .01</strong></td>
</tr>
</tbody>
</table>

Note: * *p < 0.05; ** *p < 0.01
ANOVA: Grade 2 # of Double/Concurrent Representational Tools Used

The mixed-design ANOVA executed to try to provide statistical insight on Research Question #3 (regarding Treatment and Control group ELLs’ use of double/concurrent representational tools) did not result in any such insight due to the excessive number of missing data.

Effect Size: Grade 2 # of Double/Concurrent Representational Tools Used

Table 18 summarizes the effect size data calculated for Research Question 3 regarding Grade 2 students’ simultaneous use of representational tools when answering various question types. Not as common as the findings for their Grade 1 counterparts, results from one of the five question types asked (and analyzed) of Grade 2 students appear to show a statistically significant difference between Grade 2 Treatment group students’ use of simultaneous (two) representational tools and Grade 2 Control group students’ use of simultaneous (two) representational tools: “Multiplication: Grouping” (Hedges’ g = 0.606, 95% Confidence Interval for Differences = 0.054 -- 1.166, p < .05). Similar to the findings shown in Table 17 (Total Frequency of Grade 1 students’ use of concurrent [two] representational tools), Table 20 shows that 2nd grade CGI students, again, used a greater average number of multiple, simultaneous representational tools than did their non-CGI peers on the “Multiplication: Grouping” problem analyzed, an emergent skills in second grade.
Table 18.

Total frequency of use of concurrent (two) representational tools by grade 2 treatment (CGI) and control (non-CGI) respondents for each assessment question type.

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 2 Question Type / Operation (Numbers)</th>
<th>Average # of Concurrent (Two) Representational Tools Used</th>
<th>Hedges' g</th>
<th>Independent p-value for mean difference (2-tailed t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>2.08</td>
<td>1.67</td>
<td>0.509</td>
</tr>
<tr>
<td>2</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.17</td>
<td>1.80</td>
<td>0.290</td>
</tr>
<tr>
<td>3</td>
<td>Multiplication: Grouping (4, 35)</td>
<td>1.91</td>
<td>1.30</td>
<td>0.606</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (110, 10)</td>
<td>1.92</td>
<td>1.86</td>
<td>0.052</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (25, 44)</td>
<td>1.43</td>
<td>1.63</td>
<td>0.229</td>
</tr>
</tbody>
</table>

Note: *p < 0.05.

Tables 19 and 20 provide a comparative summary of the percentage of students in Grade 1 and 2, respectively, that answered each question correctly, along with a summary of the p-value mean difference’s significance levels and the difference between the average data for the variety of representational tools used, the total number of representational tool instances observed, and the number of concurrent (2) tools used.

The Grade 1 summary table (Table 19) shows that the 1st Grade Treatment group (ELLs taught by CGI teachers) used more complex representation systems (using the # of concurrent (2) representational tool instances as a proxy) than their Control peers (ELLs taught in traditional, non-CGI classrooms). The first graders also showed this pattern in 3 out of the 5
problem types analyzed, while the 2nd grader CGI Treatment group only evidenced this finding for 1 out of 5 problem types analyzed, when compared to their non-CGI, Control group peers. The following three Grade 1 problem types where a significant difference was observed may be reflecting a curricular emphasis during Grade 1, as suggested by the Common Core and Florida Standards at the 1st Grade level: Separate/Subtraction (Result Unknown); Multiplication (Grouping); and Addition/Subtraction: Join Change Unknown (Missing Addend).

The Grade 2 summary table (Table 20) may be showing signs of floor effects, with regards to representational tool use mediated by problem types, regardless of whether the student is a CGI (Treatment) or non-CGI (Control) student. The only exception appears to be with the “Multiplication: Grouping” problem type, which is an emergent 2nd Grade content area.

Focus on Correct Grade 1 Responses (Research Questions 1, 2, and 3)

Following up on the observations made above based on the snapshots of various metrics summarized on Tables 19 and 20, which summarize observed patterns in the variety, total number of instances, and tally of concurrent (2) tools used when students solved the mathematics word problems, which the CGI students more often answered these correctly than their non-CGI peers, it naturally follows to look at these specific data when students responded the questions correctly.
Table 19.
Grade 1 Summary of correct student respondents, *p*-value mean difference significance levels, and differences of averages of the variety, total, and concurrent (2) representational tool instances.

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 1 Question Type / Operation (Numbers)</th>
<th>% of Respondents Correctly Answering (Table 2)</th>
<th>Variety of Rep. Tools (Table 4)</th>
<th>Total Rep. Tool Instances (Table 6)</th>
<th># of Concurrent (2) Rep. Tool Instances (Table 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td>p-value sig. level for mean diff.</td>
<td>p-value sig. level for mean diff.</td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>73.7% (<em>n</em> = 14)</td>
<td>61.3% (<em>n</em> = 19)</td>
<td>ns</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>Multiplication: Grouping (6, 5)</td>
<td>36.8% (<em>n</em> = 7)</td>
<td>12.9% (<em>n</em> = 4)</td>
<td>ns</td>
<td>0.23</td>
</tr>
<tr>
<td>3</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>31.6% (<em>n</em> = 6)</td>
<td>21.9% (<em>n</em> = 7)</td>
<td>ns</td>
<td>0.61</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (60, 10)</td>
<td>33.3% (<em>n</em> = 5)</td>
<td>16.7% (<em>n</em> = 4)</td>
<td>ns</td>
<td>0.07</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (15, 24)</td>
<td>36.4% (<em>n</em> = 4)</td>
<td>66.7% (<em>n</em> = 8)</td>
<td>*</td>
<td>1.37</td>
</tr>
</tbody>
</table>

Note: * *p* < 0.05; ** *p* < 0.01. ns = not significant.
Table 20.
Grade 2 Summary of correct student respondents, p-value mean difference significance levels, and differences of averages of the variety, total, and concurrent (2) representational tool instances.

<table>
<thead>
<tr>
<th>Q. Grade 2 Question Type / Operation (Numbers)</th>
<th>% of Respondents Correctly Answering (Table 3)</th>
<th>Variety of Rep. Tools (Table 5)</th>
<th>Total Rep. Tool Instances (Table 7)</th>
<th># of Concurrent (2) Rep. Tool Instances (Table 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td>p-value sig. level for mean diff.</td>
<td>p-value sig. level for mean diff.</td>
</tr>
<tr>
<td>1 Separate / Subtraction: Result Unknown (17,9)</td>
<td>84.0% (n = 21)</td>
<td>89.3% (n = 25)</td>
<td>ns</td>
<td>0.21</td>
</tr>
<tr>
<td>2 Subtraction; Compare Difference Unknown (8, 15)</td>
<td>44.0% (n = 11)</td>
<td>60.7% (n = 17)</td>
<td>ns</td>
<td>(0.01)</td>
</tr>
<tr>
<td>3 Multiplication; Grouping (4, 35)</td>
<td>36.0% (n = 9)</td>
<td>35.7% (n = 10)</td>
<td>ns</td>
<td>0.11</td>
</tr>
<tr>
<td>4 Division; Measurement (110, 10)</td>
<td>33.3% (n = 8)</td>
<td>30.8% (n = 8)</td>
<td>ns</td>
<td>0.36</td>
</tr>
<tr>
<td>5 Addition/Subtraction; Join Change Unknown (Missing Addend) (25, 44)</td>
<td>76.9% (n = 10)</td>
<td>39.1% (n = 9)</td>
<td>ns</td>
<td>(0.04)</td>
</tr>
</tbody>
</table>

Note: * p < 0.05; ** p < 0.01. ns = not significant.
When statistically analyzing the efficacy (i.e., correct responses) of varied, total, single, or concurrent representation uses by Grade 1 Control and Treatment Group students, no statistically significant differences in the variety (Research Question 1), total (Research Question 2), or average number of single representation uses (part of Research Question 3) were found.

Table 21, however, summarizes the data calculated for a second portion of Research Question 3 regarding Grade 1 students' *simultaneous* use of representational tools when *correctly* answering various question types. Results from two of the five question types asked of (and analyzed), and correctly responded by, Grade 1 students show a statistically significant difference between Grade 1 Treatment group students’ use of simultaneous (two) representational tools and Grade 1 Control group students’ use of simultaneous (two) representational tools: “Addition/Subtraction: Join Change Unknown (Missing Addend)” problem type (Hedges’ $g = 3.370$, 95% Confidence Interval = $1.565 – 5.175$, $p < .01$); and “Separate / Subtraction: Result Unknown” (Hedges’ $g = 0.833$, 95% Confidence Interval = $0.049 – 1.478$, $p < .05$). In each of these instances where a significant difference was found, the difference was characterized by Grade 1 Treatment students exhibiting a greater use of *simultaneous* use of representational tools when *correctly* answering various question types, than their Control student peers. In short, when 1st grade CGI EL students correctly answer questions involving subtraction or addition (with missing addend) problems, they use more complex representational modeling than do their 1st grade non-CGI EL peers.
Table 21.

*Total frequency of use of concurrent (two) representational tools by grade 1 treatment (CGI) and control (non-CGI) respondents correctly answering each assessment question type.*

<table>
<thead>
<tr>
<th>Q.</th>
<th>Grade 1 Question Type / Operation (Numbers)</th>
<th>Average # of Concurrent (Two) Representational Tools Used</th>
<th>Hedges' g</th>
<th>Independent p-value for mean difference (2-tailed t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td>95% Confidence Interval</td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>3.38</td>
<td>2.14</td>
<td>0.833</td>
</tr>
<tr>
<td>2</td>
<td>Multiplication: Grouping (6, 5)</td>
<td>4.17</td>
<td>3.67</td>
<td>0.138</td>
</tr>
<tr>
<td>3</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.00</td>
<td>2.00</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (60, 10)</td>
<td>1.50</td>
<td>1.00</td>
<td>0.831</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (15, 24)</td>
<td>3.00</td>
<td>1.00</td>
<td>3.370</td>
</tr>
</tbody>
</table>

Note: * p < 0.05; ** p < 0.01; # = Insufficient data available to calculate.

**Focus on Correct Grade 2 Responses (Research Questions 1, 2, and 3)**

When statistically analyzing the efficacy (i.e., correct responses) of varied, total, single, or concurrent representation uses by Grade 2 Control and Treatment Group students, one statistically significant difference in the variety (Research Question 1), and total (Research Question 2) number of representational uses were found. This statistically significant difference from each of the two analytical perspectives were found in the “Division: Measurement” type problem, respectively: variety of representational tools used (Hedges’ g = 1.096, 95% Confidence Interval = -0.007 -- 2.080, p = .046) and
total number of representational tools used (Hedges’ $g = 1.030$, 95% Confidence Interval = -0.083 -- 1.985, $p = .058$). No statistically significant differences in the average number of single representational tool uses (a part of Research Question 3) were found.

Similar to the observations noted above for Grade 1, Table 22 summarizes the data calculated for a second portion of Research Question 3: Grade 2 students’ *simultaneous* use of representational tools when *correctly* answering various question types. Results from two of the five question types asked of Grade 2 students (and analyzed), and which they answered correctly, show a statistically significant difference between Grade 2 Treatment group students’ use of (two) simultaneous representational tools and Grade 2 Control group students’ use of (two) simultaneous representational tools: “Separate / Subtraction: Result Unknown” problem type (Hedges’ $g = 0.666$, 95% Confidence Interval = -0.012 -- 1.171, $p < .05$); and “Subtraction; Compare Difference Unknown” (Hedges’ $g = 0.869$, 95% Confidence Interval = 0.015 -- 1.589, $p < .05$). In both of these instances where a significant difference was found, the difference was characterized by Grade 2 Treatment students exhibiting a greater use of *simultaneous* representational tools when *correctly* answering various question types, than their Control student peers. In short, when 2nd grade CGI EL students correctly answer questions involving various subtraction problem types (either with a result unknown or when comparing an unknown difference), they use more complex representational modeling than do their 2nd grade non-CGI EL peers.
Table 22.

Total frequency of use of (two) concurrent representational tools by grade 2 treatment (CGI) and control (non-CGI) respondents when correctly answering each assessment question type.

<table>
<thead>
<tr>
<th>Q. #</th>
<th>Grade 2 Question Type / Operation (Numbers)</th>
<th>Average # of Concurrent (Two) Representational Tools Used</th>
<th>Hedges' g</th>
<th>Independent p-value for mean difference (2-tailed t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Treatment (CGI)</td>
<td>Control (non-CGI)</td>
<td>95% Confidence Interval</td>
</tr>
<tr>
<td>1</td>
<td>Separate / Subtraction: Result Unknown (17,9)</td>
<td>2.20</td>
<td>1.63</td>
<td>0.666</td>
</tr>
<tr>
<td>2</td>
<td>Subtraction: Compare Difference Unknown (8, 15)</td>
<td>2.80</td>
<td>1.50</td>
<td>0.869</td>
</tr>
<tr>
<td>3</td>
<td>Multiplication: Grouping (4, 35)</td>
<td>1.20</td>
<td>1.67</td>
<td>0.904</td>
</tr>
<tr>
<td>4</td>
<td>Division: Measurement (110, 10)</td>
<td>2.00</td>
<td>1.00</td>
<td>#</td>
</tr>
<tr>
<td>5</td>
<td>Addition/Subtraction: Join Change Unknown (Missing Addend) (25, 44)</td>
<td>1.40</td>
<td>2.00</td>
<td>0.514</td>
</tr>
</tbody>
</table>

Note: * p < 0.05; # = Insufficient data available to calculate.

Summary of Results

The results of this study support several important, interesting, and perhaps even unsurprising findings and observations. First, they show that CGI may affect the use of complex representational systems by 1st and 2nd Grade ELs when solving a variety of arithmetic word problem types. This finding is especially evident when these CGI ELLs correctly solve these word problems. As has been seen when analyzed from other, non-representational systems-based analyses, the findings of this study show, yet again, how the particular problem type exerts
a strong influence on the use of representational systems by ELLs. The results may also have some additional relevance when the content and practices suggested be taught by the Grade 1 and Grade 2 Florida Standards and Common Core standards are considered as they are taught from a CGI-informed framework perspective. First grade ELLs in CGI classrooms tended to use more complex representational systems across a greater variety of several problem types than did non-CGI ELLs, particularly when solving these questions correctly. Second grade ELLs in CGI classrooms also used more complex representational systems than did their non-CGI ELL peers, but appeared to do so only on content emergent during the second grade (i.e. multiplication), and, again, particularly when correctly solving the word problem.

On the 1st research question investigating the variety of different representational tools used, Grade 1 ELLs in CGI classrooms used a greater variety of representational tools than their non-CGI ELL peers on a majority of problem types analyzed. This finding was not true amongst the 2nd Graders. On the 2nd research question investigating the frequency (or total number) of representational tools used, significant findings were again seen when analyzing the different problem types in both grades (rather than CGI or non-CGI classroom affiliation, which was only seen during Grade 1 CGI vs. non-CGI analyses). Again, Grades 1 and 2 CGI ELLs used significantly more total representational tools for some problem types than others, and did so significantly when compared to their non-CGI ELL peers in Grade 1.
On the ANOVA analysis of concurrent (two) representational tools used by Grade 1 and Grade 2 students, no significant differences were found either between the various question types nor between the ELL CGI/Treatment students and the non-CGI/Control ELLs, most likely because of too few data points. However, when analyzing the effect size while analyzing the independent variables (grade level, treatment or control group, and grade-specific question type) to examine how this observed dependent variable (the number of concurrent [2] representational tools used) changed when any independent variable was different while this independent variable remained constant, several significant findings were noted. Three Grade 1 question types (“Addition/Subtraction: Join Change Unknown (Missing Addend)”, “Separate / Subtraction: Result Unknown”, and “Multiplication: Grouping”) evidenced statistically significant differences between Grade 1 CGI/Treatment ELLs’ use of simultaneous (two) representational tools, when compared to Grade 1 non-CGI/Control ELLs’ use of simultaneous (two) representational tools. An additional analysis when the problem type was solved correctly showed that two of these three problem types (“Addition/Subtraction: Join Change Unknown (Missing Addend)” and “Separate / Subtraction: Result Unknown”) were solved by Treatment vs. Control ELLs using concurrent (2) representational tools in statistically significant ways. A similar statistically significant result was found when comparing Grade 2 ELLs’ (CGI/Treatment vs. non-CGI/Control) use of simultaneous (two) representational tools when solving “Multiplication: Grouping”
(which was also one of the Grade 1 problem types, with lower magnitude numbers, also found to be statistically significant in this regard).

Results of mixed design two-way Analyses of Variance (ANOVAs) provided a span of results for Grade 1 and for Grade 2 for each of the three main research questions of this study. Specifically, Table 23 below summarizes the question type and treatment findings by research question and grade that were found to be statistically significant.

Several general observations, in closing, from the snapshots provided above in Tables 21 and 22 are worth noting. The treatment (CGI) students correctly answered the assessment questions given to their respective grades more often than the control (non-CGI) peers did. This was true in all but one Grade 1 question (question #5, “Addition/Subtraction: Join Change Unknown (Missing Addend) (15, 24)), and in the 2 (presumably easiest Questions #1 and #2) out of the 5 Grade 2 questions asked and analyzed. Only 1 of the 5 mean differences in the variety of Grade 1 representational tools used were significant (for question #5), and only 1 of the 5 mean differences in the total Grade 2 representational tool instances (question #4), so the harder questions asked and analyzed in each grade. Furthermore, 3 of the 5 mean differences in the concurrent (2) Grade 1 representational tool uses, and 1 of the 5 mean differences in the concurrent (2) Grade 2 representational tool uses were significant. Focusing on the column listing the differences in the respective variables’ averages shows that these significant differences were almost-always noted when CGI students used a
Table 23.

Summary of specific statistically significant Question Type and Treatment findings, by research question and grade.

<table>
<thead>
<tr>
<th>Research Q. #</th>
<th>Significant</th>
<th>Grade 1</th>
<th>Grade 2</th>
</tr>
</thead>
</table>
| 1: "VARIETY of Representational Tools Used…" | ANOVA | • Gr.1&2/Q.1 > Gr.1/Q.4  
• Gr.1/Q.5 > Gr.1/Q.4  
• CGI ELLs > non-CGI ELLs (when answering Gr.1/Q.5)  
• CGI ELLs Only: Gr.1/Q.5 > Gr.1/Q.4 | ns | ns |
| | Effect Size | • CGI ELLs > non-CGI ELLs (when answering Gr.1/Q.5) | | |
| | Efficacy | ns | CGI ELLs > non-CGI ELLs, when correctly answering Gr.2/Q.4 |
| 2: “TOTAL # of Representational Instances used…” | ANOVA | • CGI ELLs > non-CGI ELLs (when answering Gr.1/Q.5)  
• CGI ELLs Only: Gr.1/Q.5 > Gr.1/Q.3 | ns | ns |
| | Efficacy | ns | CGI ELLs > non-CGI ELLs, when correctly answering Gr.2/Q.4 |
| 3: “COMPLEXITY (Total # OF Concurrent (2)) Representational Tools Used…” | Effect Size | CGI ELLs > non-CGI ELLs, when answering:  
• Gr.1&2/Q.1; Gr.1/Q.2; Gr.1/Q.5 | CGI ELLs > non-CGI ELLs, when answering Gr.2/Q.3 |
| | Efficacy | • Gr.1&2/Q.1; Gr.1/Q.5 | CGI ELLs > non-CGI ELLs, when correctly answering:  
• Gr.1&2/Q.1; Gr.2/Q.2 |
| Statistically Significant Question Type Key (By Grade Level): | | | |
| | Gr.1&2/Q.1: “Separate / Subtraction: Result Unknown” (17,9) | Gr.2/Q.3: “Multiplication: Grouping” (4, 35) |
| | Gr.1/Q.2: “Multiplication: Grouping” (6,5) | Gr.2/Q.4: “Division: Measurement” (110, 10) |
| | Gr.1/Q.3: Subtraction; Compare Difference Unknown” (8, 15) | Gr.2/Q.5: “Addition/Subtraction: Join Change Unknown (Missing Addend)” (25, 44) |
| | Gr.1/Q.4: “Division: Measurement” (60, 10) | |
greater variety of, a greater total number of, and a greater number of concurrent (2) tool uses than did their non-CGI same grade peers. Putting it all together, in turn, the data seem to support the observation that CGI students engaged more varied, total, and simultaneous representational tools when correctly solving a variety of mathematics word problems, than did their same-grade non-CGI peers.
CHAPTER 5

DISCUSSION

Summary of Findings

This study was conducted to try to establish baseline patterns regarding how ELLs enrolled in two different 1st and 2nd grade mathematics classroom settings (one led by a teacher familiar with the CGI teaching and learning framework, and the other led by a teacher not familiar with these pedagogical and assessment approaches) use a variety of, a total number of, and a coordinated, simultaneous use of various representational tools when solving relatively typical grade-level mathematics word problems. Results of mixed design two-way Analyses of Variance (ANOVAs), as well as the effect size upon comparisons of dependent and independent variables, provided mixed results supporting insights into both the assessment questions posed to the students, as well as comparisons between the two treatment student groups.

With regards to the variety of representational tools used (Research Question #1), the only aspect of the study found to be significant was the variety of representational tools used by Grade 1 CGI/Treatment students when solving particular types of questions, as they used a greater number of different representational tools when solving “Addition/Subtraction: Join Change Unknown (Missing Addend)” problem types than when they solved “Subtraction; Compare Difference Unknown” problem types. Furthermore, a statistically significant
difference was found when comparing the (greater) variety of representational tools used by CGI/Treatment ELLs solving “Addition/Subtraction: Join Change Unknown (Missing Addend)” problem types to the (less) variety of representational tools used by their non-CGI/Control peers when solving the same problem.

Regarding the total number of representational tools used (Research Question #2), the statistically significant findings are again seen between question types in Grade 1 ("Subtraction: Compare Difference Unknown" and "Addition/Subtraction: Join Change Unknown (Missing Addend)" question types) and Grade 2 ("Separate/Subtraction: Result Unknown", "Subtraction: Compare Difference Unknown", and "Multiplication: Grouping" question types), but not between the total representational tool uses of Treatment ELLs vs. Control ELLs. These findings also show that, when solving a presumably more difficult Grade 2 question type ("Multiplication: Grouping"), ELLs used more total representational tools than when solving presumably easier questions ("Separate/Subtraction: Result Unknown" and "Subtraction: Compare Difference Unknown").

The third Research Question investigating the use of concurrent (2) representational tools was more difficult to analyze statistically because of far fewer overall instances noted to establish any definite tool use patterns, but an analysis of effect sizes between independent and dependent variables still resulted in a statistically significant findings that were question-type dependent. Three Grade 1 question types ("Addition/Subtraction: Join Change Unknown (Missing Addend)", "Separate / Subtraction: Result Unknown", and
“Multiplication: Grouping”) were solved by Grade 1 Treatment and Grade 1 Control ELLs using a total, different number of concurrent (2) representational tools that was statistically significant.

Conclusions

Relationship of Representational Tool Use to Problem Types

The findings of this study build upon prior studies’ findings, doing so from a representational tool use perspective: When the unit of analysis is the use of multiple representational tools while problem solving, the problem types being solved become a key factor consideration. This study agrees with Secada (1991) that problem type variations (such as those typical in CGI classrooms and on the assessment used in this study) make it possible for a greater variety of all students to successfully participate in solving them. Furthermore, the breadth and scope of the extensive CGI literature point to representations and representational tool use as being at the heart of problem selection, strategy selection, solution sharing, and problem solvers’ comparison of various solution strategies. This study suggests that ELLs in individualized assessment settings transfer to (and apply during) individualized assessment settings, the well-established connections between representations, representation tool use, and problem types that teachers familiar with the CGI framework have modeled for
decades while guiding their mathematics students towards developing mathematics comprehension and proficiency.

The findings indicating that the overall effects of CGI on representational tool use in varied problem types is more pronounced during Grade 1 than Grade 2 supports Carpenter, Fennema, Franke, et al.’s (1999) findings that the direct modeling strategies involving tangible, visible representational tools and their uses are replaced with more abstract counting strategies and number facts. These findings also support previous work that showed that students (such as those in classrooms taught by teachers implementing the CGI framework) who have frequent opportunities to use representational tools as they solve word problems, provide a way for others to access their mathematical thinking (e.g., Fennell & Rowan, 2001; Harries & Barmby, 2006; Kamii, Rummelsburg, & Kari, 2005; Kendrick & McKay, 2004).

**Effects of CGI Instruction during Individualized Assessments**

A review of state Mathematics standards for 1st and 2nd graders shows that Multiplication problems are not emphasized in the elementary school curriculum until 3rd grade (Florida State University, 2014). Therefore, the significant findings showing that the 1st and 2nd grade CGI/Treatment ELLs used a greater variety of representational tools in more complex ways (than their non-CGI/Control peers) when solving “Multiplication; Grouping” speaks to how CGI teachers are preparing their ELLs (and all their students) to effectively solve
arithmetic word problems years before they are formally expected to do so, and
years before their traditional instruction, non-CGI peers do so. Although there is
a general introduction to multiplication in the Grade 2 standards Cluster 3
(“MAFS.2.OA.3: Work with equal groups of objects to gain foundations for
multiplication,” Florida State University, 2014), the standards suggest educators
focus on multiplication beginning in Grade 3 as part of three clusters of
Operations and Algebraic Thinking subdomain (“MAFS.3.OA.1: Represent and
solve problems involving multiplication and division”, “MAFS.3.OA.2: Understand
properties of multiplication and the relationship between multiplication and
division”, and “MAFS.3.OA.3: Work with equal groups of objects to gain
foundations for multiplication”). The multiplication focus suggested by the
standards beginning in Grade 3 tends to support approaching multiplication from
a geometric arrays perspective, while CGI supports a repeated addition approach
(as early as Grade 1) as a basis to the evolution of grouping, essential to
success in multiplication problem solving. The effects of CGI instruction on
individualized assessments of early elementary school students, therefore, would
appear to be that these CGI students become more familiar earlier with some
mathematical concepts and problem solving approaches (such as multiplication)
that their non-CGI peers don’t develop until later in their elementary years.

As was noted in Chapter 2, the CGI classroom is characterized by a
number of features that are evident in individualized assessments settings as
well as absent from those. The findings of this study allude to earlier work
(Heritage & Niemi, 2006) in documenting students’ abilities to transfer some of
the features from their CGI classrooms to independent assessment settings. Given the quantitative nature of this study and the fact that the principal unit of analysis was the frequency of representational tool use, when the CGI students display a greater variety of representational tools used as well as a greater number of total representational tools used, this finding maps directly back onto CGI teachers’ ongoing efforts to increase the frequency and variety of solution strategies, a function of the total number and different number of representational tools used during their individualized assessments.

The fact that a significant difference of total and diverse representational tools used was observed when comparing ELLs from CGI classrooms to their non-CGI peers as well as when analyzing this representational tool used for particular problem types in both grades 1 and 2, reflects another important characteristic of the typical CGI classroom. CGI students that participated in this study appeared to be more selective in which problem types yielded significantly different representational tool uses. While it is just speculation at this point, future studies that build on this current study may be able to indicate whether the greater number and variety of representational tools used were engaged as part of a corroborating effort to have multiple solution strategies lead to the same solution, or perhaps these represent a second or even third strategy implemented to correct a solution that the student felt was incorrect (but didn’t know for certain), or whether this difference in number and variety of tools used was a “Hail Mary” approach where the student tried everything to see if any of the attempted solution strategies and their results “felt” right. Students that
engaged in any of these activities may have done so after having experienced some of these reasons to engage in more and different representations-rich solution strategies.

Implications for the Assessment of ELLs

The findings of this study provide some support for continuing to use word problems that implicitly encourage the use of a variety of representational modeling tools to help ELLs solve those problems correctly. Because most assessments do not take into account the product and process of using representational modeling tools, this study provides some evidence that these are pertinent when problems are solved correctly. Further qualitative analysis is necessary to document how ELLs use these representational modeling tools when a problem is solved incorrectly. However, it is important to note that the very concept of correct and incorrect responses by ELLs solving arithmetic word problems may need to be expanded to include consideration for the representational models that ELLs create, to give these some weight when holistically considering the efficacy of a response. The practice of students showing their mathematics “work” has always been in place. For ELLs, however (but for all students potentially), giving credit for correct representational modeling shown in a problem’s solution that may contradict a final “wrong” answer could increase the achievement levels of these students while also targeting areas in need of additional instruction and scaffolding. This would be
particularly true where the requirement to provide a complete sentence is necessary in order for the problem to be considered completely correct. The opportunities to use representational modeling as a way to bridge ELLs inherent understanding of a problem with their misunderstanding certainly lends itself to consideration when assessing their proficiency and comprehension levels. Finally, the focus on representational modeling when assessing ELLs would help reduce yet another discrepancy between the classroom instructional setting and the individualized assessment test center – a move that might encourage ELLs to show the same proficiency during a test as they did while in class.

The findings of this study may have additional implications on the accommodations that are provided to ELLs during individualized assessments. Typically states and organizations (for example, New York State Education Department, 2019), allow for accommodations for ELLs including extended time, testing in a separate location, bilingual dictionaries (providing one-to-one translations of words, without including definitions or explanations), and/or glossaries. This study’s findings appear to suggest that other accommodations, such as representational systems, may be helpful to students solving mathematics word problems and should be considered as future accommodations for ELLs during individualized assessments.

**Implications for the Instruction of ELLs**

As with implications for the assessment of ELLs, the findings of this study support the continued use and consideration of representational modeling in
classrooms with ELLs. Particularly relevant is the findings from Research Question #3 regarding the use of complex, concurrent representational modeling tools. As Ainsworth (2006) and others have long promoted, the use of multiple representational modeling tools under the guide of a teacher (who intrinsically believes ELLs have a number of assets that they bring to the mathematics classroom) can serve to supplement, complement, and generally reinforce the meaning-making that ELLs are trying to communicate, but may be unable to successfully do so with text and oral discussions only. The absence of specific details regarding the English proficiency levels of the ELLs who partook in this study makes it premature to state (but invites the thought of) whether ELLs of lower language proficiency may benefit more than their more proficient peers when given and encouraged to use multiple representational tools during assessment settings.

**Implications for CGI Professional Development**

The CGI Professional Development framework appeared to be relevant throughout this study. In turn, the findings of this study are consistent with the continued use of representational modeling tools, discussion regarding their use (both amongst educators and students), and perhaps serve as a reminder that ELLs (and students in general) may model (in both the sense of building as well as enacting) what is modeled for them. There are transfers of practice and concepts to the individualized assessment setting that ELLs enrolled in CGI
classrooms appeared to display through this study. The focus on representational modeling tools as an ubiquitous proxy for these practices and concepts that exists in both class and testing environments in the absence of the social and collaborative peer and teacher supports, is one that (it is hoped) this study provides some evidence for continuing to emphasize and perhaps expand.

**Limitations**

**Possible Limits to CGI based on Differences Between Instructional and Assessment Settings**

This study tests whether ELLs enrolled in CGI classrooms use representational tools and systems in ways noticeably different from how ELLs who are enrolled in more traditional math classrooms use representational tools. In the event that no differences had been found between how students in different classrooms use representational systems, this study may have identified a potential limitation in the CGI professional development framework: the CGI-informed classroom instruction and practices that include modeling, justification, and critiquing others’ solutions and representations, may not have consistent analogues during individualized assessments. The option exists, then, to acknowledge and accept this limitation, or to acknowledge this limitation and then revisit current CGI professional development training to see if any changes might help address this observation. Updated CGI training may more explicitly and purposefully include in-class representational modeling activities, tools, and
practices that will become more familiar to students before they are assessed later.

Despite (or, perhaps, because of) the many affordances of CGI that have been discussed and whose effects were noted in some of the findings above, this study faced some inherent limitations. As noted, the CGI framework is holistic and classroom-based, rather than prescriptive and inclusive of formal individualized assessments such as those used for this study. CGI teachers observe the goals and learning objectives of CGI but are not expected to follow a particular set of pedagogical processes or fixed problem sets. As such, attempts to establish direct analogues between what a CGI teacher might implement or practice in his/her classroom and those that a student might exhibit evidence of during an individualized assessment setting, are bound to be somewhat subjective, presumptive, and limited, particularly without the qualitative inclusion of student interviews which could corroborate some of the findings observed.

Inherent in this limitation, of course, is the fact that the CGI framework, problem types, and solution strategies have always operated exclusively within the context of the academically social, dynamic, teacher-present classroom, and was never intended to be analyzed during individualized assessment settings. When the absence of a teacher during the individualized assessment of interest (who can provide the kind of probing and guiding questioning, adaptive problem selection, and purposeful groupings of classmates) is coupled with the absence of fellow CGI students (who can provide alternate strategies and representational tool use for comparison and modeling of potentially more mature solution
pathways), the change in setting from a CGI classroom to a more restricted and less interactive assessment setting is certain to add a number of limitations that might affect the validity and reliability of the findings of this study.

Similarly, while CGI classrooms are replete with numerous representational tools that are available to students, and several of these were known to have been the same ones that were provided to the participants during the video recorded assessments analyzed for this study, there is no certainty that the set of tools available during the assessment for the particular problem types asked were exactly those available to students when they might have practiced on these problem types before. In fact, none of the students participating in this study were asked whether they were familiar with all the representational tools provided, and even if they had been, there is no guarantee that their claimed familiarity with these equated to also knowing the benefits and drawbacks of using (and when to use) each representational tool. Also, because CGI teachers do not prescribe or tell students which representations to use and why, preferring instead to allow students to discover these aspects for themselves, the actual representational tool used has not been the focus of prior CGI research.

Limitations of the Study Design and Assessment Tool Design

There are a number of limitations to this study primarily based on it having been a derivative research effort off a larger study with far different research foci (e.g., teachers vs, students, framework implementation vs. representational tool...
use, no purposeful case matching by language proficiency). Some of the limitations include the size and relevant characteristics of the pool of participants, the absence of certain pre-study participants’ data, and the resulting limitations to analytical methodologies.

From many perspectives, the larger CGI study (that served as the source from which the participants for this study were selected) was a very large research study involving hundreds of students and scores of teachers, as well as numerous researchers and support personnel. One of the greatest limitations to this study was the inability of the research team collecting the overall video data corpus to obtain the discrete English proficiency levels of the ELLs who were interviewed for the studies from the school districts. As noted earlier, while the WIDA Consortium provides clear guidance on which of six different tiers an ELL might be classified under, the ELLs in this study where considered to be one monolithic group of students.

When the goals of this current study were detailed, and then the students and video data that would meet the language proficiency and CGI/non-CGI classroom enrollment criteria were all said and done, the final subset from the large overall data set was so greatly diminished as to almost be insufficient for this proposed study. In fact, initially, it was the intent of this study to compare the representational tool use of CGI and non-CGI ELLs and EP students. However, in addition to the absence of consistent case-matching between the 4 cohorts (CGI ELLs, CGI EPs, non-CGI ELLs, and non-CGI EPs), the overall number discrepancies between the hoped-for cohorts was simply too large to overcome
for this study. Hence the decision was made to just focus on ELLs, as the relevance to how they might use representational tools to bridge the language and context of the word problems and their solution strategies might be of some interest to the larger educational research community.

Once the subset group that would be analyzed for this study was finalized as ELLs enrolled in classrooms taught by teachers implementing the CGI framework in their classrooms (and their non-CGI peers), the matter of sufficient data sources again became a concern when the research questions (particularly #3) were finalized. Although there was sufficient (but not ample) data sources available for the first two research questions regarding the variety of different tools used and the total number of representational tools used, there was some concern about whether there would be a sufficient number of data points related to the third research question (regarding simultaneous use of 2 representational tools during problem solving). As this information was not known a priori, the study (and data coding) proceeded as if there would be no lack of data issues. When the data were analyzed for this final research question, however, the initial concerns proved to generally hold true and there was an absence of sufficient data to make definitive statements using analysis of variance methodology; an alternate statistical analysis was used but was inconsistent with the analyses applied to the other two research questions.

As mentioned earlier, the use of SPSS to analyze the ANOVAs for this study was an unexpected limitation, given how the software handles null responses within a set of otherwise complete data by listwise deleting these from
the overall analysis. As a result of this known shortcoming in the SPSS software, the total $n$ analyzed for several of the variables in this study may have been lower than they would have been if an alternate analytical software (like R, for example; R Core Team, 2014) or alternative analyses (such as Linear Mixed Effects Equations, for example; Bates, Mächler, Bolker, & Walker, 2014) had been used.

Although the focus on concurrent use of representational tools was anticipated to be a potential data desert, it was pursued as a research focus anyway because the existing literature regarding how students coordinate the use of multiple representational tools, and why and how they use these multiple representational tools, may be of great interest in enhancing the existing work in CGI solution strategies (which, to date, has not explicitly included representational tool use).

On the matter of the assessment itself, this study focused on analyzing a subset of the dozens of arithmetic questions that were asked, and to focus on the representational tool use during the five word problem questions that were either identical or relatively similar for both grade 1 and grade 2 assessments. This means that the other question formats (Number Facts and Solving Equations, and Equations: True/False and Multidigit Computation) were not analyzed from a representational tool use for this study. It should be noted that the various representational tools were provided for some, but not all, of these other question format sections before and after the word problem section analyzed for this study. A limitation here, therefore, is simply that, for each of the five-word
problem types analyzed for this study, there was only one example of each problem type upon which to base the significant (or not significant) findings on. Clearly, an additional example of each word problem type might have been helpful in supporting the overall findings of this study. Since students were not surveyed about whether and how well they knew about the affordances and limitations of each of the representational tools provided, there was an unfounded assumption that the provisioning of the representational tools was an addition to their solution strategies “toolkit”, without knowing if the tools provided might, in fact, have been detractors and distractors from an otherwise focused set of solution strategies that did not involve the provided representational tools.

Finally, the Goldin and Kaput (1996)/Goldin (1998) representational systems taxonomy used for this study may benefit from an update that reflects the proliferation of technology in the classroom and in students’ lives since they originally proposed their taxonomy. External representations carried out in technological and/or virtual platforms that may not easily fall into Goldin and Kaput’s 5-level representational systems taxonomy: *imagistic*, *formal notational*, *planning/executive*, *verbal/syntactic*, and *affective* (not used for this study). For this study, and based on research in the area of gestures and embodied mathematics and feedback from fellow researchers, I made a slight modification to Goldin and Kaput’s nomenclature, distinguishing fingers from being considered to be part of the same *imagistic* category as Base-10 cubes or Unifix/Snap cubes. A similar update may need to be explored in future work in this field if this
taxonomy is still considered to be the most applicable, notwithstanding any shortcomings in how to account for technological representations.

**Recommendations for Future Research**

This study lies at the beginning stages of a research area that has many tangential relevancies to CGI, ELLs, and overall use of representational tools in mathematics problem solving, but have not been as explicit in trying to understand this phenomenon in assessment settings with the ELL population targeted for this study engaging in problem types such as those present in this study's assessment tool. Therefore, there are many, many directions where future researchers may explore further the initial findings of this study. These suggested areas for future research might be best be classified into two major areas: units of analyses and number and nature of participants.

This study was a quantitative study focused on a relatively aggregated unit of analysis: the tally of instances of representational tools used while independently solving the word problems in the assessment. This area of study is poised to benefit from qualitative studies that begin to explore the underlying nature of these tallies. Future research may want to analyze which tools were used to see if a particular representational tool used was more common for a particular problem type (so do join and separate problems induce students to use snap cubes and base-10 cubes over base-10 rods and fingers because the former are designed to be unitary representations that can easily be joined
together or separated apart, while the latter are, by design, non- or semi-unitary representational tools). Further, subsequent research might analyze whether particular representational tools were used in a particular order, or during a particular phase of problem solving (do fingers get used initially, while verbal self-narrations typically serve to close the problem-solving episode, for example). This study began setting up some important future work that should be expanded upon regarding the efficacy of the use of representational tools when correctly (or incorrectly) solving word problems. Further, it might be an important contribution to study whether students from CGI classrooms generated multiple solution strategies to reach the same conclusion, or whether the solution was reached by using only one solution strategy. A slightly modified version of this study that might address similar prior work (e.g., Branch, 2000; Fonteyn, Kuipers, and Grobe, 1993) might include an obligatory, explanatory debriefing phase after every problem is solved or attempted so that students’ explanations of their problem solving process can be compared to their actual evident representational tool use and resulting models – a practice that at least the CGI students should be quite accustomed to. Future qualitative studies might also seek to develop or use coding rubrics to analyze students’ generated representations, and not just the representational tools used to effect these. Of particular interest might be to compare representations for the same problem type and trace back to see if a particular characteristic of the students is more likely to result in a particular type of representational model. Schleppegrell’s (2007) work on the mathematics register and Fuson, Smith, and LoCicero’s
(1997) work on the unitary and sequence-and-tens counting concepts (described in Chapter 2) remind us that, since this study also did not establish any one-to-one mapping of the quantities, context, or mathematical concepts explicitly or implicitly found in the problem type to the resulting representational models, future researchers may wish to establish additional relationships between the mathematics register and the representational tools used by taking on research trajectories gauging how much of what is given or what is induced through a problem’s text makes it to their final representational model of the problem’s solution strategy. A far more intricate study might be to compare these individualized assessment explanations of problem-solving pathways a student might have taken to how the same student explains problem solutions while in class. It may also be necessary to step back and provide some baseline work on which representational tools are used when ELLs (and/or EPs) solve the other types of arithmetic problems that were asked on this assessment: the prior Number Facts and Solving Equations section, and the True/False and Multidigit Computation section.

Most traditional instructional approaches and assessments seldom include explicit instruction on visual representations. Some prior work, however, has evaluated curricular emphases on visual representations that have had a measurable effect on student assessment outcomes. For example, the University of Chicago School Mathematics Project’s elementary curriculum Everyday Mathematics (EM) trains teachers to use manipulatives and other visual representations to scaffold students’ thinking while solving mathematics
story problems. Researchers have found that EM students using representational modeling tools (such as geoboards, tape measures, and pattern-block geometry templates) achieved higher scores on area and perimeter questions (Fuson, Carrol, & Drueck, 2000) than students who did not use these representation tools. These results lend some support to curricula like EM that train teachers to facilitate students' visual, concrete exploration of more typically abstract concepts in geometry and data measurement. Where future studies may provide some key insights is in trying to disaggregate if a correctly-solved assessment problem is correctly solved due, in some evident part, to a proficiency in the representational systems related to the language and concepts comprising and undergirding the question. This study contributes to that research trajectory by helping establish if ELLs (in classrooms led by CGI-familiar and non-CGI familiar classroom teachers) use representational models and tools in specific ways for specific problem types.

Future studies may contribute to the literature claiming that schematic representational modeling reduces the linguistic complexity of math word problems by helping ELLs make meaning of the texts that they read (Martiniello, 2009). These follow-up studies could investigate whether students generate schematic representational models when solving word problems with less-complex language (capturing, for example, if a student used representational tools to schematically model a quantity and/or action called for in a word problem, or if the resulting representational system was more non-schematic and deviated considerably from the problem posed).
This study provides some data basis for future researchers to evaluate whether the use (or not) of representational modeling tools by ELLs when solving mathematics word problems was more common when ELLs solved word problems correctly. Building on work in this area already initiated (Avalos, Medina, & Secada, 2015) regarding how prior math proficiency may, not only predict ELLs proficiency at post-test points, but also which interventions for ELLs may be taken up (as problem solving entry points) and which might not, this study suggests continuing that work from a representational systems approach. This study also provides researchers and mathematics educators with data regarding whether ELLs respond to (or sidestep) representational modeling and the tools used to generate them when solving particular types of arithmetic word problems. This guidance may help educators and other assessment designers and publishers know whether purposefully integrating these models and tools into assessments of ELLs’ mathematics word problem solving is worth further exploration.

Notwithstanding an ensuing qualitative study that focuses on ELLs’ word problem solving achievement, this study supports educators who explicitly discuss, teach about, and assess the representational modeling of word problems (and their student-generated solutions) in conjunction with the language and mathematical concepts already being taught. When educators provide explicit classroom instruction (or even just the time and space for discussion) on representational modeling in word problems, they provide students with greater familiarity with assessments traditionally discussed from a
purely language-based perspective. Because these pedagogical emphases do not preclude EP students from also activating representational modeling tools to communicate and make sense of the mathematics on which they are also assessed, the outcomes of this study could benefit all learners.

This study provides quantitative data showing a difference in representational modeling and tool use patterns by CGI and non-CGI students as well as by certain problem types that set the stage for future research into the possible threats to an assessment’s validity discussed in Chapter 1. Future studies may focus on interpreting how the language and mathematical concepts assessed by a given problem are reflected in students’ representational modeling. Furthermore, future researchers may focus far more on the analysis of the problem-solving process rather than simply on the right vs. wrong efficacy of the student’s responses. Additionally, future studies could investigate whether students in similar CGI learning classrooms evidence representational modeling practices during assessment settings and/or on particular problems that are notably different from their non-CGI peers.

An important area poised for future research in this topic is that of the composition of the participants of those studies. This study focused on macro-level demographic factors of CGI vs. non-CGI students, and on whether they were 1st or 2nd graders, and whether or not they were designated as ELLs by their school district. Of these three variables, the question regarding the precise proficiency level of ELLs was unexplored as a result of the absence of WIDA consortium level data regarding details of which of the six-tiered English
proficiency levels the ELLs that participated in this study characterized the ELLs. Future researchers would serve ELLs and the education research community well to determine and consider the level of English proficiency of ELLs they study in the future since all ELLs are not the same. Since this study focused only on the representational tool use while ELLs were solving word problems, there is a comparison to EP students that this study did not pursue. This might satisfy Abedi, Hofstetter, & Lord’s (2004) suggestion to compare the performance of EPs to ELLs with and without the availability of representational tools, just to ascertain how the representational modeling tool accommodations might be affecting students’ achievement. If future studies are conducted with students in higher grades (middle or high school students, for example), the gender factor of participants might play an important mediating role. This study did not explicitly account for pre-established mathematics achievement proficiencies prior to analyzing participants representational tool use on the word problems posed in this study. However, it is reasonable to ask (and therefore systematically investigate) whether students from different achievement and proficiency levels use representational tools in different ways (or at all).
References


Schoen, R. C., Secada, W., & Tazaz, A. M. (2015). *Results after the first year of a randomized controlled Trial of CGI*. Presentation during 2015 annual CGI meeting.


Appendix A – Representational Systems’ interactions with Counting

Previous studies (unrelated to CGI) have used representational modeling to inform other aspects of pedagogical content knowledge by detailing some of the most frequently observed counting concepts and strategies evidenced by young students: the unitary, decade-and-ones, and sequence-tens-and-ones conception. Students manifest these counting conceptions through representational modeling (Fuson, Smith, & Lo Cicero, 1997). In modeling a unitary conception, for example, students count the number of unifix cubes present in natural number order (so 34 cubes would be counted as “1, 2, 3, 4, 5,…32, 33, 34”). Practice and time eventually brings students to model the decades-and-ones conception by assigning meaning to the individual digits of numbers (less than 100) as either a decade or as extra ones beyond a particular decade, so that 34 would mean that the 3 represents the thirty decade and the 4 represents four. This conception, originally identified by Murray and Olivier (1989) and discussed in other studies (Fuson, 1990; Fuson, Smith, & Lo Cicero, 1997; Fuson, Wearne, et al., 1997), is particularly relevant because it helps explain why some children will write or otherwise model thirty-four as the number “304” upon hearing “thirty four”.

The sequence-tens-and-ones counting concept extends the decades conception to account for the number of single tens units formed by the decade/“tens place” part of the quantity, or the 3 tens units in 34. While engaged in the sequence-tens-and-ones conceptual counting level, a student is counting
by tens (so for 34, “10, 20, 30”) while keeping track of (and not anticipating) the count of three. This conceptual count sets the stage for the separate-tens-and-ones conceptual counting, where a student regards the decades/"tens place" as and the ones as two separate units, counting each unit by ones: 1, 2, 3 tens and 1, 2, 3, 4 ones. The importance to the proposed study of being aware of these three potential counting concepts, is that representational modeling and the engagement of corresponding modeling tools may all be a function of which of the three counting concepts (unitary, decades-and-ones, or sequence-tens-and-ones) the student models while problem solving.

Researchers have observed that some younger students focus on mathematics words (which more effectively supports sequence-tens concepts), while some of their peers focus on the written numerals (more closely aligned with developing separate-tens and separate-tens-and-ones concepts) (Fuson, Smith, & Lo Cicero, 1997, pp. 742-743). Researchers noted the prevalence of different counting concepts depending on the class' language of instruction. English-speaking class students more often exhibited separate-tens-and-ones counting concepts, and showed more proficiency at using grouping or adding to explicitly create additional tens or when breaking a ten. The Spanish-speaking class students, on the other hand, appeared to prefer sequence-tens counting concepts and used the ten-sticks and dots drawing method developed for this study and base-ten blocks provided to avoid making additional tens when
counting by tens and then by ones in order to reach the next decade/"tens place" value. (Fuson, Smith, & Lo Cicero, 1997, p. 760).

Furthermore, researchers noted that while the students in the Spanish-speaking class more often counted by tens and then by ones (10, 20, 30, 31...34), the vast majority of English-speaking class students counted the tens and ones (1, 2, 3 tens and 1, 2, 4 ones) (Fuson, Smith, & Lo Cicero, 1997, p. 758). While noting different counting methods used between the two different groups of students, the study’s authors described the differences in the students’ representational modeling as they “enclosed or opened quantities to make a ten differently, and they made recordings to denote making another ten in different ways” (Fuson, Smith, & Lo Cicero, 1997, p. 762).

Some studies have shown that when home languages are considered as a part of a research agenda to address mathematics instruction for ELLs, that there are some positive observations to be made that may be evidenced by representational modeling activities. For example, Fuson et al.’s (1997) analysis of the structure of Spanish words for two-digit numbers revealed that, while Spanish uses a list of “decade words (diez, veinte, treinta, cuarenta, cincuenta, sesenta, setenta, ochenta, noventa)” (p. 744) and that the relationship between these and the words for the first nine natural numbers (one, two, three, etc.) is not clear, that this observation holds true for the parallel English words such as “twenty” and “two”, “thirty” and “three”, and “fifty” and “five.” Even a quick perusal
at the Spanish decade words for 60 ("sesenta") and 70 ("setenta") indicate a but-
one-letter similarity between their spellings and pronunciations that might
confuse Spanish students when using their home language, while potentially
benefitting their understanding based on a more-than-one-letter and one–syllable
distinction between the English version of these decade words ("sixty", and
"seventy"). Similar to English, Spanish words for numbers greater than 20
naturally lend themselves to the decade-and-ones conception (where each
number word and each digit takes on a meaning as either a decade or as the
additional ones), an observation that can then be incorporated into mathematics
instruction and representational modeling of these numbers. For students from
European and East Asian cultures, these constructions hold, so that those
students (and their ELL teachers) are more fully served by focusing on the
“unitary” (counting two digit quantities by ones) and “separate tens-and-ones”
(where the number of tens units is counted separately from the number of ones
units) (Fuson et al., 1997, p. 741) counting concepts that may differ from English
in other languages.

Previous work with CGI have highlighted three related, but distinct,
representational modeling strategies each representing increasing levels of
abstraction: direct modeling using tens rods, combining the abstraction of the first
quantity with the counting of the second quantity, and using fingers to count on
by tens (Carpenter et al., 1996). Carpenter et al. (1996) note that the fingers in
the third strategy (counting on by tens using fingers) took on a role different from
the blocks or cubes in the first two strategies: instead of the fingers being substitutes for the blocks or cubes, “the counting sequence itself had become an object of reflection, and as such, it could be counted” (p. 12). If students’ evident algorithms are, therefore, generated because of cumulative abstraction of prior concrete representational modeling practices, then it stands that verbalized narrations of what students modeled earlier are also a form of a solution that exists without the need to position representational modeling tools as explicit referents (Carpenter, Fennema, & Franke, 1996). These authors propose that the very representational models which uses those tools as well as the applicable numbers and quantities themselves become objects of reflection, so that students are eventually able to manipulate numbers and quantities without needing to also manipulate representational tools (Carpenter, Fennema, & Franke, 1996).

When students use and manipulate representational modeling tools to represent their counts and thinking, there is a logistics problem that presents itself for a teacher related to permanence of this evidence for later analysis. Fuson, Smith, and Lo Cicero (1997) noted that base-ten blocks and other physical representational modeling tools are used for each problem as they are needed and leave no record after a problem is solved or a class period has ended. The teacher, therefore, is unable to analyze or react to a particular or a series of representational model(s) once the next problem’s model has replaced them or the representational tools have been returned for storage. In the cited
study, therefore, the authors introduced a system using vertical lines as ten-sticks and dots as units that could permanently record quantities. These drawn representational models represent a base-ten structure whose decade/"tens place" column depictions may give more insight into students’ generating and counting at each ten that was not as clearly evident with students’ more typical, less structured unit drawings (Fuson, Smith, and Lo Cicero, 1997, p. 747). Furthermore, the use of a recording protocol that relies on pencil/pen/marker and paper begins to weaken students’ reliance on physical representational modeling tools to solve mathematics problems. This approach shapes the proposed study by motivating the use of a mechanism by which all representational models and the use of the corresponding tools would be documented for subsequent analysis and comparison despite the student resetting and using these tools for the next word problem.
Appendix B – Instance Codes and Labels – Definitions and Examples

The following is provided as a reference tool to help explain the particular details regarding the coding of students’ use of representational modeling tools.

The free video-coding software used for this coding is BORIS (Behavioral Observation Research Interactive Software). BORIS is available for Microsoft- Windows, Mac OS X, and GNU/Linux platforms at http://www.boris.unito.it. The Python3 source code is released under the GNU General Public License at https://github.com/olivierfriard/BORIS.

(The video-coding software [StudioCode] used during an earlier pilot study that preceded this study ceased operations and therefore was not used.)

General Notes:

- The start-point should be marked at the beginning of the audible discourse or visible action that contains the instance(s) of interest to be coded. Once an instance code is recognized, the coder should mark a start-point that provides just enough context for that instance Code to be understood if excerpted out of the context of the larger cognitive video interview.

- The end-point should be marked at the end of the audible discourse or visible action that contains the instance(s) of interest to be coded. Once coding has begun on an instance code(s), the audiovideo footage should play until an end-point that includes the completion of the audible discourse or visible action and does not include an evidently different
context, aspect of discourse, or model-based action. The end-point should be such that a viewer that watches or listens to the entire excerpt feels they have listened to or seen a complete segment that starts just prior to the instance of interest, and runs through until just after this instance is concluded. Furthermore, a listener/viewer should not feel that additional information about the closing aspects of the audiovideo clip are necessary, if the end-point was appropriately set.

- It is common that two instances of the same Code may occur within close time proximity of each other, or even concurrently. The BORIS software makes it possible to code these sequential instances to fractions of a second of each other. If a clear start- and end-point have been established, and the next instance’s start-point is within seconds of the prior instance’s end-point, these should still be considered separate and distinct codes. This is suggested because it allows for a more accurate Frequency count of Instances – an essential data element that can answer related research questions. Fortunately, these data are relatively easy to document in the BORIS software. In general, two clearly distinct iterations of an instance Code should not be coded as one instance.

- Instance codes are to be applied anywhere that there is a public, audible interaction and/or visible action. In cases where a student may be whispering to him/herself but audibly enough to be discerned, these self-talk instances should be coded. Given this study’s focus on representational models and representational modeling tools, actions
involving representational modeling tools (Base-10 Cubes/Rods/Flats, Markers/Writing on Paper, Fingers, and Unifix/Snap Cubes) should be coded anytime the video data provide clear and verifiable evidence that the tool(s) was/were used. However, if a student, for example, counts using their fingers but under a table where the fingers are not visible to the video camera’s field of vision, or if the student is “mouthing” something but only a professional lip-reader might be able to determine with certainty what s/he said, then these instances shall not be coded, as they lack the level of certainty that should support every coded instance.

- There is no minimum or maximum time limit per instance code, so long as the general notes above and the specific start- and end-points below hold for each instance Code duration.

- The beginning of one instance Code does not eliminate or otherwise alter the beginning of an additional, different instance Code, even if the first instance Code has not ended. Similarly, there is no requirement nor expectation that an instance Code that begins first should end first. The resulting overlaps may be fully embedded instance codes or they may only partially overlap. These embedded, concurrent instance Codes are reflective of the multi-layered nature of multiple models and modeling being engaged by students in multiple contexts during mathematics assessment settings.
“Base-10 Blocks” Coding

The “Base-10 Cubes”, “Base-10 Rod”, and “Base-10 Flat” codes are three of seven representational modeling tools whose use was provided and/or permitted while students were solving the assessment problems. These codes are activated any time and for the entire visible period during which the student purposefully and/or methodologically touched, moved, assembled, disassembled, or counted any of the (sky blue-colored) base-10 blocks available (i.e., 1-count unit cube, 10-count rod, or 100-count flat). These manipulatives are relatively small, monochrome (sky blue), 3-dimensional quadrilateral figure with rectilinear corners, and they do not interlock.

Examples when “Base-10 Cubes”, “Base-10 Rod”, and “Base-10 Flat” are used (and therefore coded) may include:

(1) to model a number provided or calculated during the reading or solving of an assessment problem;

(2) dynamic models/modeling of addition, subtraction, or grouping problems (by engaging, joining, or grouping additional cubes, rods, and/or flats or separating/removing any of these that might have been initially engaged;

(3) using these base-10 cubes, rods, or flats in conjunction with snap cubes to model different numbers
Figure 1. Video excerpts of examples of students’ use of Base-10 Blocks.

<table>
<thead>
<tr>
<th>Base-10 Blocks Ex. 1</th>
<th>Base-10 Blocks Ex. 2</th>
<th>Base-10 Blocks Ex. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instance # 07</td>
<td>Instance # 08</td>
<td>Instance # 01</td>
</tr>
<tr>
<td>Time w/in Instance: 00:07.80</td>
<td>Time w/in Instance: 00:25.76</td>
<td>Time w/in Instance: 00:09.86</td>
</tr>
</tbody>
</table>

15 seconds of relevant audio discourse including above freeze-frame (30-second span):

None | None | None

“Markers/Writing on Paper” Coding

The “Markers/Writing on Paper” code is activated any time and for the entire visible period during which the student purposefully and methodologically reached for and uncapped the provided marker(s) to then draw or write anything on the assessment questions paper provided.

Examples when “Markers/Writing on Paper” are used (and therefore coded) may include:

1. a student reaching for a marker and/or paper to prepare to write or draw something related to the problem being assessed;

2. when the marker is already in the student’s hand, the coding for this instance begins when the student either uncaps a covered marker or gestures to use an uncovered marker to write or draw on the assessment problem paper;
(3) ending this coding when the student finishes drawing or writing with the marker on the paper (and there is no further use of the marker or paper for at least 10 seconds), and/or when the student covers or caps the marker.

Figure 2. Video excerpts of examples of students’ use of Marker/Paper.

<table>
<thead>
<tr>
<th>Marker/Paper Ex. 1</th>
<th>Marker/Paper Ex. 2 &amp; 1 Tick = 1 Unit Ex. 1</th>
<th>Marker/Paper Ex. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instance # 01</td>
<td>Instance # 02</td>
<td>Instance # 03</td>
</tr>
<tr>
<td>Time w/in Instance: 00:10.56</td>
<td>Time w/in Instance: 01:02.23</td>
<td>Time w/in Instance: 01:00.66</td>
</tr>
</tbody>
</table>

15 seconds of relevant audio discourse including above freeze-frame (30-second span):

None

Interviewer: “So your answer is what?”
Student: “Two.”
Interviewer: “Two?”

“Fingers” Coding

The “Fingers” code is activated any time and for the entire visible period during which the student purposefully and methodologically moves their fingers (most often from a closed fist with retracted fingers, to an open palm with extended finger and/or fingers). This instance is typically evident in response to a numerical model provided or created while solving a problem.

Examples when “Fingers” are used (and therefore coded) in evident kinesthetic ways may include the following:
(1) Often (but not always) accompanied by a verbal enumeration or counting on or counting down; and/or

(2) when using (typically) one finger to touch a representational modeling tool (i.e., base-10 blocks or snap cubes) while counting it/them.

Figure 3. Video excerpts of examples of students’ use of Fingers.

<table>
<thead>
<tr>
<th>Fingers Ex. 1</th>
<th>Fingers Ex. 2</th>
<th>Fingers Ex. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instance # 25</td>
<td>Instance # 26</td>
<td>Instance # 01</td>
</tr>
<tr>
<td>Time w/in Instance: 00:01.60</td>
<td>Time w/in Instance: 00:04.76</td>
<td>Time w/in Instance: 00:05.73</td>
</tr>
</tbody>
</table>

15 seconds of relevant audio discourse including above freeze-frame (30-second span):

| None | None | “Four, five, six, seven…thirty-one, thirty-two, thirty-three…” |

“Snap Cubes” Coding

Like the “Base-10 Cube/Rod/Flat”, the “Snap Cubes” code is activated any time and for the entire visible period during which the student purposefully and methodologically touches, moves, assembles, disassembles, or counts any of the available multi-colored Snap Cubes. These representational modeling tools are larger than Base-10 blocks, vary in colors though each is single-colored, have rounded corners, an interlocking mechanism that allows these Snap Cubes
to be combined upon snapping interlocking pieces of one onto the other, and allows for the separation of these Snap Cubes via the separation of these same interlocking components.

Examples when “Snap Cubes” are used (and therefore coded) in evident kinesthetic ways may include the following:

(1) to model a number provided or calculated during the reading or solving of an assessment problem;

(2) to model addition, subtraction, or grouping problems (by joining, separating, or grouping additional cubes);

(3) using these snap cubes in conjunction with base-10 cubes, rods, or flats to represent different numbers, objects, or operations.

Figure 4. Video excerpts of examples of students’ use of Snap Cubes.
move them because they’re so tall. And I know what I can make out of these."

Interviewer: “OK, so how many Legos does he have now?”
Student: “How many Legos?”
Interviewer: “How many Legos does he have now?”
Student: “56.”
Interviewer: “So he only has 56?”
Student: “I can make a castle out of these… This took a long time.”

“Verbal Response” Coding

This instance code is applicable any time an audible utterance and/or and audible response was provided to an assessment problem. This Code is used to note when a student generates a verbal, audible narration, response, or answer to a problem (instead of a written or gestured response, for example). This code is most often used during the “Solve Phase” of an assessment’s problem solving episode, so they would be seen anytime between the moment a problem is presented and/or read to a student and culminate at the moment a response is finalized to a problem. This code is, perhaps, the easiest way to capture, and then in turn, analyze the answers to assessment problems spoken by students.

Example when “Verbal Response” are evident (and therefore coded) may include the following: verbally narrating a counting on of Base-10 Cubes (i.e., “one, two,
three, four, etc.), an audible self-narration of a process (“I think I will connect these five snap cubes because…”), or verbally stating “four” after a student writes the number 4 as the solution to a problem.

“Verbal Explanation” Coding

This instance Code captures the generation of student explanations during cognitive think alouds when students are prompted to explain, clarify, or specify a solution pathway or step after a problem’s solution has been finalized. Therefore, unlike the “Verbal Response” code, it would only be coded during the post-solution “Explain Phase”. This code is activated immediately following an interviewer’s prompt if the student provides an explanation or other clarification. These typically begin with statements like “Because…”, “I did…”, “I noticed that…”, or “Since I saw that…”

Figure 5. Video excerpt transcript of examples of students’ Verbal Explanations.

<table>
<thead>
<tr>
<th>Verbal Explanation Provided (in response to prompt) Ex. 1</th>
<th>Verbal Explanation Provided (in response to prompt) Ex. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video: 3225-2-Moore014</td>
<td>Video: 3126-2-Tapp014</td>
</tr>
<tr>
<td>Instance # 01</td>
<td>Instance # 02</td>
</tr>
</tbody>
</table>

**Relevant audio discourse:**

Interviewer: “You counted up from what number? What number did you start at?”
Student: “9”
Interviewer: “9? And then you counted up to 13?”:
Student: “Mmm-hmm”

While responding to: WP/Gr2/Q09: Alicia went to the beach with her family for 2 days. She found 28 seashells on Saturday and 43 seashells on Sunday. How many seashells did she find?

Student: She had 10, 20, so I had 20, and I added 8 to it.
Interviewer: So I saw you had… you had your two piles, right?
Interviewer: OK. Did you use your fingers or did you do it in your head?" Student: "My fingers."

Student: Mmm-hmm. Interviewer: OK. One was 28 and this pile was 43. And then what did you do? Student: I put these together so I can count them, like, the both of them. Interviewer: OK. Student: And then I added these. Interviewer: and so how did you count these? Can you count again out loud? Student: 10, 20, 30, 40, 50, 60, then I counted... Interviewer: ... by ones. OK. Thank you. I saw exactly what you did. Thank you

“Right” Solution Labeling

This label is activated towards the end of each problem solved, as this is when it is more certain if a student has solved a problem correctly. This label is intended to capture whether a student’s final offer of a solution to a given problem is mathematically correct. Note that the intent is to capture the final answer offered by the student, as this would be the one most likely to be graded if there were no interviewer nor video-record of the assessment available, and only a paper document or online submission form were available.

“Wrong” Solution Labeling

This label is intended to capture whether a student’s final offer of a solution to a given problem is mathematically incorrect. Note that the intent, again, is to capture the final answer offered by the student, as this would be the one most likely to be graded if there were no interviewer nor video-record of the assessment available, and only a paper document or online submission form were available. If a “Wrong” Solution label is applied, the coder should use
BORIS’ “comment” feature to free-form type exactly what the Wrong answer offered was. This is also the opportunity to provide any additional commentary that may be of interest to researchers, including and particularly, if a verbalized final, incorrect response is actually different from a correct response, as evident in the representational modeling available.

“Skipped” Labeling

This label should be activated if the student contributes no discernible solution whatsoever and/or if s/he hesitates to try to solve a problem and shows no signs of progressing towards any solution of the problem. This code is often used after a student is asked if they would like to move on to the next problem before they gave an answer.

“Not Asked” Labeling

This label should be activated if the interviewer (or other non-student factor) leads to a situation where a student is never shown, read, or asked to solve a particular question. The interviewer may make this decision if s/he applies the “mercy” rule described in the study where no additional questions are asked if a student has incorrectly answered two questions in a row. It may also be the case, however, that a student falling ill or a building fire alarm going off during the assessment leads to an early end to the entire assessment, and therefore, means a student never even saw a problem(s).