Characterizing and Assessing Covariational Reasoning in Introductory Physics Contexts

Charlotte Marshall Zimmerman

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:
Suzanne White Brahmia, Chair
Peter Shaffer
Jason Detweiler

Program Authorized to Offer Degree:
Physics

University of Washington

Abstract

Characterizing and Assessing Covariational Reasoning in Introductory Physics Contexts

Charlotte Marshall Zimmerman

Chair of the Supervisory Committee:
Suzanne White Brahmia
Department of Physics

Quantitative literacy—the use of mathematics to describe and understand the world—is an essential skill. Physics has an opportunity to contribute to how students develop quantitative literacy, as much of physics curriculum centers on making sense of quantitative models. One facet of quantitative literacy in physics is *covariational reasoning*: how changes in one quantity affect changes in another, related quantity. Covariational reasoning is at the heart of developing and making sense of quantitative models, and is central to graphical reasoning. Much of the research on covariation lies in mathematics education; the language of covariation has only recently begun to be used amongst physics education research. Research in physics and mathematics education has demonstrated that reasoning mathematically in physics contexts is distinct from reasoning mathematically in a context-free way. Early indications suggest that, similarly, covariational reasoning is likely different in physics contexts than in mathematics. Moreover, research into quantitative literacy more generally has shown that it is unlikely to improve in physics classrooms without direct instruction. Therefore, there is a need to characterize and understand *physics* covariational reasoning towards developing instructional activities that can be used in physics classrooms to help students develop quantitative literacy.

The research presented in this dissertation represents a collection of work that provides a foundation for instructional activity development. We describe research into characterizing

and operationalizing physics covariational reasoning through a series of studies that examine how physics experts reasoned while generating graphical models. The results of this study, together with prior research in the field, are organized into a framework of covariational reasoning: the Covariational Reasoning in Physics (CoRP) framework. We present this framework in this dissertation, and describe how it can be used towards identifying learning outcomes for introductory physics courses and beyond, identifying proto-expert resources that students may already have when entering physics courses, and developing instructional interventions that attend to improving students' quantitative literacy. We then present two short reflections on two assessment tools, the Physics Inventory of Quantitative Literacy (PIQL) and the Generalized Equation-based Reasoning inventory of Quantity and Negativity (GERQN), that are designed to measure physics quantitative literacy across a wide range of student populations. These assessment tools can be used to measure the impact of instruction on students' physics quantitative literacy, and thus are a necessary tool towards designing activities that are supported by research. This dissertation concludes with a reflection on how these pieces can be used together for future steps towards the development of instructional materials.

TABLE OF CONTENTS

Page

List of Figures iii

List of Tables viii

Glossary x

Chapter 1: Introduction 1

Chapter 2: Literary Review 8
2.1 Mathematical Reasoning in Math Class is not the same as Mathematical Reasoning in Physics Class 8
2.2 Reasoning about Symbols and Quantities in Physics Contexts 11
2.3 Covariational Reasoning in Mathematics and Physics Education 13
2.4 Assessing Quantitative Literacy: the PIQL 16

Chapter 3: Characterizing Expert Physics Covariational Reasoning in Graphing Tasks 18
3.1 Introduction 18
3.2 Background 20
3.3 Methods 23
3.4 Findings 33
3.5 Discussion 49
3.6 Conclusion 54

Chapter 4: Empirical Evidence for the Blended Nature of Physics and Mathematics 57
4.1 Introduction 57
4.2 Background 58
4.3 Methods 61
4.4 Findings 66

4.5 Discussion 74
4.6 Conclusion 77

Chapter 5: The Covariational Reasoning in Physics Framework 79
5.1 Introduction 79
5.2 Background 80
5.3 Characterizing Covariational Reasoning in Physics 83
5.4 Relevance of the covariational reasoning framework 102
5.5 Conclusion 109

Chapter 6: Assessing Covariational Reasoning with the Physics Inventory for Quantitative Literacy (PIQL) and the Generalized Equation-Based Inventory for Quantity and Negativity (GERQN) 112
6.1 Introduction 112
6.2 Exploring student facility with "goes like" reasoning in introductory physics 113
6.3 Assessing physics quantitative literacy in algebra-based physics: lessons learned 124

Chapter 7: Towards Intervention Development 136
7.1 Finalizing the Validation of the GERQN 137
7.2 Validity Beyond the University of Washington 139
7.3 Characterizing PQL Across the Major 140
7.4 Identifying Learning Outcomes and Incoming Student Resources 140
7.5 Designing Instructional Interventions 141
7.6 Assessment and Iterative Review 142

Chapter 8: Conclusion 143

Bibliography 147

Appendix A: Expert Covariational Reasoning Codebooks 163

Appendix B: Math-Physics Blend Timelines 166

LIST OF FIGURES

Figure Number Page

1.1 (A) A distribution of PIQL (Physics Inventory for Quantitative Literacy) scores from administration in the introductory physics sequence at the University of Washington. These data are an updated version of our previously published results [166], and represent results from 801 students who enrolled in all three courses. "Pre-Mech" refers to their scores coming into the university before physics instruction, "Post-Mech" refers to scores at the beginning of their second quarter in physics (Introductory E&M), and "Post-EM" refers to scores at the beginning of their third quarter in physics, after taking Introductory E&M. The data are modeled with normal distributions. (B) A comparison of the normalized change (how much a student's score changed in relation to the amount of points they had to gain or lose) with their incoming PIQL score. 2

2.1 The graph used in Bajarcharya and Thompson's study of student quantitative reasoning about "area-under-the-curve" problems [11]. 9

2.2 (A) the measurement symbolic form [42], (B) the quantity symbolic form [160] 11

2.3 Two popular modeling cycles in mathematics and physics education research; (A) by Blum and Leiß[16] and (B) by Uhden, Karam, Pietrocola, and Pospiech [148] 12

2.4 An example item from the PIQL [166, 1]. 16

2.5 The correlation between incoming PIQL score (before any physics instruction at the University of Washington) and final course grade after one quarter of introductory physics. These data represent a sample of 1024 students. The linear regression shown is described by: (Course Grade) $= (60 \pm 1) + (0.39 \pm 0.02)$(PIQL Score). $R^2 = 0.31 \pm 0.03$. 17

3.1 Stills from the animations associated with tasks administered in Study 1: (A) Going Around Tacoma, (B) Ferris Wheel, and (C) Square Track. All tasks asked participants to create a graph that related a particular quantity (distance to Tacoma, height of the cart, and distance from the cart to the wall, respectively) to the total distance traveled of an object in constant motion. These graphing tasks are identical those in the original study conducted by Hobson and Moore [65], with the exception that we changed the place names in Going Around Tacoma. 25

3.2 Stills from the animations associated with the first task administered in Study 2. Participants either saw (A) Electric Charge or (B) Gravitation. The tasks prompt participants to create a graph that relates either the electric or gravitational potential and the total distance traveled of the probe or spaceship, as it moves at constant speed from start to finish. 26

3.3 Stills from the animations associated with the second and third tasks administered in Study 2: (A) Drone, in which participants were prompted to create a graph that relates the height to the angle of the drone, flying in the arc shown, and (B) Intensity, in which participants were prompted to create a graph that relates the intensity measured by the probe at the side of the liquid column to the total distance traveled of the probe as it moves from the bottom of the column to the top. 26

3.4 An example time on task chart for a faculty member from Study 3. The horizontal axis represents normalized time, or the fraction of time on task. The symbol represents the time the category was coded, and the band represents the duration that the interviewee was engaged with that kind of reasoning. . 29

3.5 The modes identified in our data are representative of a combination of covariational reasoning mental actions as described by mathematics education research and the reasoning devices we identified in Section 3.4.2. 42

3.6 Four examples of the interplay between Function and Mental Actions across Studies 2 and 3. These examples were chosen because they are illustrative of the trends we saw for both faculty and graduate students. The horizontal axis represents time on task, and is normalized across participants. 43

3.7 This figure depicts the modes we observed experts engage in when modeling, and the interplay between the mathematics mental actions and reasoning devices for each mode. The modes are shown to lie along a spectrum from Applying to Generating Models. 43

3.8 A graduate student's work from Ferris Wheel that demonstrates Function Generating reasoning; the point and line segment they used are highlighted in green. 48

3.9 The time participants in Study 2 and 3 spent Function Knowing, Function Choosing, or Function Generating as a percentage of time spent engaged in function reasoning. 49

4.1 Czocher's redraft of Blum and Leiß's modeling cycle [16, 37]. The original Blum and Leißcycle is reproduced in Fig. 2.3. 59

4.2 Stills from the animations associated with (A) Electric Charge and (B) Gravitation. The tasks prompt participants to create a graph that relates either the electric or gravitational potential and the total distance traveled of the probe or spaceship, as it moves at constant speed from start to finish. This figure is reproduced from Chapter 3 for ease of the reader. 62

4.3 Stills from the animations associated with (A) Drone, in which participants were prompted to create a graph that relates the height to the angle of the drone, flying in the arc shown, and (B) Intensity, in which participants were prompted to create a graph that relates the intensity measured by the probe at the side of the liquid column to the total distance traveled of the probe as it moves from the bottom of the column to the top. This figure is reproduced from Chapter 3 for ease of the reader. 62

4.4 A visual representation of when participant statements were assigned the "Quantity," "Task Context," and "Models" code categories. The length of the bars corresponds to the time the participant was engaged in that line of reasoning. The horizontal axis represents time, normalized to the length of each interview. Therefore, it can be thought of as a percentage of their total time on task. The outlined regions are where "Quantity" appears without "Task Context." Gray vertical lines represent when participants switched tasks. 67

4.5 Percentage of time spent reasoning in a blended way, which we measure as Quantity and Task Context coded together, abstractly, which we call "mathematical riffing" and is measured as Quantity coded without Task Context, and other reasoning, which is measured as when Quantity is coded without Task Context, but the transcript reveals that the participants are not riffing. 69

4.6 Ernest's written work for the Drone item. 72

4.7 Percentage of time spent reasoning: in a blended way, as measured by Task Context and Quantity coded together; mathematically riffing, as measured by when Quantity is coded without Task Context; and Other, as measured by when Quantity is coded without Task Context but the transcript reveals that the reasoning is not fully aligned with our definition of mathematical riffing. Percentages were calculated as percent time on task for each participant, and averaged by population to produce the charts shown. 73

5.1 A still from a covariational reasoning graphing task that asks experts to create a graph relating the distance traveled by a Ferris wheel cart and the height of the cart. 84

5.2 Top: Task provided here to exemplify use of the Expert Behavior reasoning devices. For this task, create a graph of the potential energy of the car-Earth system as a function of distance traveled for a toy car of mass m. The loop portion of the track is estimated as circular. Bottom: Annotated qualitatively correct graph for task. 98

5.3 Bottle-filling task figure 101

6.1 The Flag of Bhutan. The prompt for this item asks students to **select all of the following quantities** that are larger by a factor of 1.5 when the length and width of the flag are both increased by a factor of 1.5: (a) The distance around the edge of the flag, (b) the amount of cloth needed to make the flag, (c) the length of the curve forming the dragon's backbone, (d) the diagonal of the flag, and (e) none of these. Students are prompted to choose all answer choices that apply. We believe the correct answers to be (a), (c) and (d). . . 119

6.2 Ferris Wheel. The prompt for this item asks students to identify which expression correctly identifies how h, the height of the cart, directly changes with s, the distance traveled by the rider, where the radius of the Ferris wheel is given by R_0: (a) $h(s) = \sqrt{s^2 + R_0^2}$, (b) $h(s) = R_0 \exp(s/R_0)$, (c) $h(s) = R_0 - R_0 \cos(s/R_0)$, (d) $h(s) = s^2/(2\pi R_0)$ 121

6.3 A PIQL item that has been adjusted for the GERQN due to physics context. (A) shows the original item, and (B) shows the revised item for the GERQN. 128

6.4 A PIQL item that has been adjusted for the GERQN due to mathematical level. (A) shows the original item, and (B) a linear version on the current iteration of the GERQN. 130

6.5 Student responses to the PIQL and GERQN versions of Spherical/Cylindrical Bottle and Growth items. The PIQL data were collected from a calculus-based introductory mechanics course (N = 240) and the GERQN data from an algebra-based introductory mechanics course (N = 225) at the same large, R1 university in the pacific northwest. 131

6.6 The GERQN item "Growth," where students are asked to compare the growth of two children using a graph of growth rate. The correct answer is B, as Alex has a higher growth rate during the entire year. 132

7.1 Early statistics from the 2022-2023 administration of the GERQN. These data represent a matched set of 185 students, meaning every student enrolled in the full introductory, algebra-based physics sequence. Difficulty is a measure of % correct (e.g. a difficulty of 0.8 on an item means that 80% of students chose the correct answer for that item). Discrimination is a measure of how correlated the score on that item is with the score on the overall test. 139

B.1 A visual representation of when participant statements were assigned the "Quantity," "Task Context," and "Models" code categories. The length of the bars corresponds to the time the participant was engaged in that line of reasoning. The horizontal axis represents time, normalized to the length of each interview. Therefore, it can be thought of as a percentage of their total time on task. The outlined regions are where "Quantity" appears without "Task Context." 167

LIST OF TABLES

Table Number Page

2.1 A summary of the covariational reasoning mental actions (MA) used in this study. The framework we provide here is a slightly modified version of that summarized by Jones in 2022 [28, 141, 76]. We use names and descriptions from Jones' summary. Numeric labels and example behaviors were added by the authors and correspond to the original mental actions from the 2002 framework, with the addition of 1.5 for "precoordination", for ease of reference within this paper. 15

3.1 A summary of the differences between the populations and tasks across the three studies. 23

3.2 The reasoning devices identified from the physics experts' descriptions of their reasoning as they thought out loud while solving novel graphing tasks. Each reasoning device is associated with—but not the same as—a covariational reasoning mental action identified by mathematics education research (described in Table 2.1). 36

3.3 The ways in which experts engaged with mathematical modeling while using covariational reasoning during graphing tasks. 45

4.1 Codes associated with the code categories Quantity, Task Context, and Models. 64

5.1 A summary of the covariational reasoning mental actions (MA) frameworks developed by mathematics education researchers [**?**, **?**]. Summary adapted from Jones [76]. This table is repeated for ease of the reader. 82

5.2 Current version of a framework to describe the use of covariational reasoning in physics modeling. 87

6.1 Examples of ways in which question statements were rewritten on the GERQN to be more uniform and easily parsed by non-native English speakers. 134

A.1 Study 1 code categories that emerged during data analysis. 164

A.2 Studies 2 and 3 code categories that emerged during data analysis. These codes were developed using thematic analysis [97] such that they describe emergent patterns from the data. While an effort was made to set aside the codebook from Study 1, many similar themes arose. We connect these code categories to those of Study 1 in the descriptions for the ease of the reader. . 165

GLOSSARY

CONCEPTUAL BLEND: A conceptual blend is the subconcious, interwoven meaning of two ideas. This analysis was popularized by linguistics Fauconnier and Turner and is used in physics education to refer to the inseparability of mathematical and physics ideas [52].

COVARIATIONAL REASONING: How one reasons about how changes in one quantity affect changes in another, related quantity.

GERQN: The Generalized Equation-based Reasoning inventory of Quantity and Negativity is a version of the PIQL for algebra-based physics classrooms. It is pronounced "gherkin."

PQL: Physics Quantitative Literacy, which is defined by White Brahmia as "Quantitative literacy in the context of physics, has the potential to be an important learning outcome for all students taking introductory physics" [166].

PIQL: The Physics Inventory for Quantitative Literacy is an assessment tool for measuring PQL, validated at the University of Washington for calculus-based physics students and beyond. It is pronounced "pickle."

PROCEPTUAL UNDERSTANDING: A combination of the words "*proce*dural" and "con*ceptual*", proceptual understanding denotes reasoning that is informed both by making sense of the procedures involved and the concepts behind why those procedures make sense [56].

QUANTITATIVE LITERACY: The ways in which one uses familiar mathematics to make sense of the world in sophisticated ways [139, 102, 166].

ZONE OF PROXIMAL DEVELOPMENT: Defined by Vygotsky as "The distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peer" [156].

ACKNOWLEDGMENTS

It is difficult to express how grateful I am to my collaborators for all their support—of myself and the work presented here—throughout my graduate career. I am deeply appreciative of my advisor, Suzanne White Brahmia, and close collaborator, Alexis Olsho. It has been a privilege to have had the opportunity to learn from and work with these thoughtful, creative, empathetic, and impressive women.

Suzanne, the joy and care you bring to your research group and collaborations has modeled for me what a career in academia can be: a place of collective, creative learning. Participating in the research spaces you have grown has been a strong example of teaching and fostering teamwork, and a huge factor in my decision to pursue academia. My words here cannot do justice to the long hours spent encouraging me, generating bubble and arrow diagrams, late-night editing, and providing academic, professional, and emotional support. Your mentorship has changed how I approach the world; you have made me a stronger collaborator, more thoughtful researcher, and kinder person.

Alexis, your advice, patience, and thoughtful approach to research have deeply shaped this work, and formed a foundation of support for me throughout graduate school. Your care towards research and writing has taught me how to develop an argument, and your empathetic ear encouraged me through so many periods of doubt. The hours we spent debating over what "goes like" *really* means are uncountable—thank you for sticking with me. I know you insist we are colleagues, and you were not my mentor—regardless, your impact has encouraged me to grow as a person, and I am immensely grateful.

I would also like to thank the PIQL collaboration—Suzanne White Brahmia, Andrew Boudreaux, Trevor Smith, Philip Eaton, Alexis Olsho, Qirui Guo, and Ella Henry. I have

regularly benefited from all of your wisdom, expertise, and sense of humor. Andrew, your approach to your students and colleagues with consistent care and mindfulness has made me a better teacher and collaborator. Trevor and Philip, your quantitative genius, and extreme patience, has been a highlight of my graduate experience. I am deeply appreciative to have colleagues with whom I can contribute to creative and exciting research while laughing and feeling a part of a strong community.

I have also worked closely in collaborations with both Michael Loverude and Mieke de Cock during my graduate career. To Michael: thank you for your patience and guidance. I have learned so much from you about research design, data analysis, and writing. Your mentorship has deeply informed how I strive to show up as a researcher, and our chats have helped me become a more compassionate person. To Mieke: thank you for your support and for providing me with so many opportunities to develop my skills as a researcher. Working with you has informed how I think about quantitative research design and helped shape my perspective of mentoring.

My thesis work would not have been possible without the support of the PER community, and all of those who have come and gone through the Physics Education Group at the University of Washington. To Peter and Paula: thank you for your collaboration and support. Peter, learning from you in and out of the classroom has had a huge impact on how I strive to show up as a teacher. Particular thank you's to Anne Alesandrini, Lisa Goodhew, Dean Bretland, Jared Canright, Jesse Ashworth, Kristin Kellar, John Goldak, Yasmene Elhady, Taylor Gurreithun, Al Snow, and Qirui Guo for your advice, friendship, and encouragement. There are too many members of the Physics Education Research community and in the Physics Department at the University of Washington who have had significant impacts on me as I developed my research identity to name here—you know who you are. Thank you all for your warmth, inspiration, and support.

To Marty Baylor, without whom I never would have persevered in the physics program

at Carleton College nor ever considered myself a “physics person.” Your mentorship taught me how to learn: thank you.

The support of my friends and family outside of physics have been indispensable. I’m deeply grateful for my family: Mom, Dad, and Grace. Your encouragement and sense of humor have helped me persevere and remember what is important. To Erika: you made it possible for me to graduate, and I am so grateful to have gained a friend in the process. To Omar: thank you for laughing and crying with me, and for keeping me grounded. To Eliza, Jeff, Mariya, Brandon, Nick, Ryanne, John Ferre, Jared, Josh, John Goldak, Martha, and many others: your humor, joy, and encouragement has kept me going. Thank you. You all have made me a better friend, teacher, and colleague.

This material is based upon work supported by National Science Foundation Grants No. DUE-1832836, DUE- 1832880, DUE-1833050, DGE-1762114, DUE-2214765, and DUE-2214283. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

DEDICATION

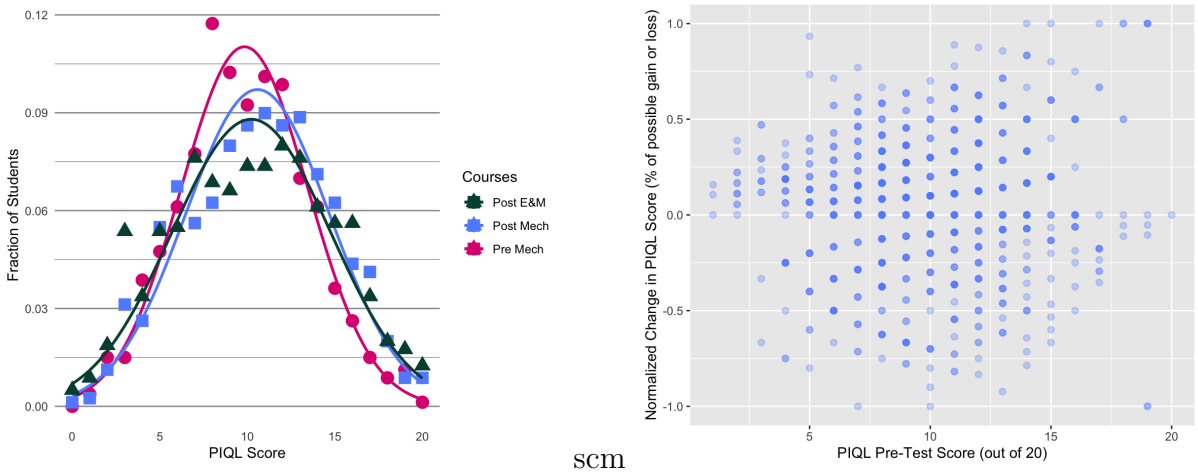
To my family, who first fostered my curiosity, taught me never to take things too seriously, and always asked "Why x?"

Chapter 1

INTRODUCTION

Quantitative literacy—the ways in which we use mathematics to interpret and make sense of the world around us—is an essential part of day to day life. As modeling real-world observations quantitatively is at the heart of physics, physics courses are excellently suited to help students develop their quantitative literacy for future careers in STEM and beyond. In addition, it has been well-established in mathematics and physics education research that reasoning in pure math contexts is fundamentally different from reasoning in physics contexts, or contexts in which the quantities are situated in the real world [130, 13, 84, 67, 116, 163]. We cannot simply assume that the mathematics used in physics courses "comes along" with mathematics instruction. Rather, it must be integrated into the ways physics instructors teach their students to reason quantitatively.

In physics contexts, we refer to this reasoning as Physics Quantitative Literacy (PQL). White Brahmia, Olsho, Smith, Broudreaux, Eaton and Zimmerman recently demonstrated that PQL is unlikely to improve over the course of an introductory physics sequence, even in classrooms that are using highly refined, research-based instructional materials designed to improve conceptual physics reasoning [166]. This research was situated at the University of Washington; the researchers administered an assessment tool, called the Physics Inventory for Quantitative Literacy (PIQL), to measure PQL across the introductory physics sequence over the course of several years. An updated version of their results is shown in Figure 1.1. Research in physics education has used Cohen's d effect size to measure differences between pre- and post-instruction; an effect size of about $d = 1$ is typical on conceptual physics assessments [96]. The effect size on the data shown in Figure 1.1 is much lower: $d = 0.2$ from PreMech to PostMech and $d = 0.1$ from PostMech to PostE&M. From these data,



scm

(A) (B)

Figure 1.1: (A) A distribution of PIQL (Physics Inventory for Quantitative Literacy) scores from administration in the introductory physics sequence at the University of Washington. These data are an updated version of our previously published results [166], and represent results from 801 students who enrolled in all three courses. "Pre-Mech" refers to their scores coming into the university before physics instruction, "Post-Mech" refers to scores at the beginning of their second quarter in physics (Introductory E&M), and "Post-EM" refers to scores at the beginning of their third quarter in physics, after taking Introductory E&M. The data are modeled with normal distributions. (B) A comparison of the normalized change (how much a student's score changed in relation to the amount of points they had to gain or lose) with their incoming PIQL score.

we conclude that students' PIQL scores do not change significantly over the three course sequence. We consider this evidence that many students are completing an introductory physics sequence with little to no improvement in their quantitative literacy, even when instruction is thoughtfully designed and aligned with current research practices.

To examine the change in PIQL score more closely, Figure 1.1 includes a comparison of the normalized change in PIQL score to students' incoming PIQL scores ("PreMech"). Normalized change is a measure recommended in physics education research as it compares how much a score has changed to how much it could have changed. It is related to normalized gain, and defined as:

$$c = \begin{cases} \frac{\text{post}-\text{pre}}{100-\text{pre}} & \text{if post} > \text{pre} \\ \text{drop} & \text{if post} = \text{pre} = \text{max or } 0 \\ 0 & \text{if post} = \text{pre} \\ \frac{\text{post}-\text{pre}}{\text{pre}} & \text{if post} < \text{pre} \end{cases} \tag{1.1}$$

At first glance, this plot may appear to show a sweeping trend upward for those with initially high pre-test scores (shown in the upper right) and a sweeping trend downward for those with initially low pre-test scores (shown in the lower left). However, it is important to note that normalized change becomes increasingly binned at extreme ends of the spectrum—students who initially scored a 19 can only have a normalized change of 1, if they improve on the assessment. Therefore, we suggest this pattern is a feature of normalized change and the data demonstrate little improvement and no obvious patterns as to which students may be benefiting from instruction.

Preliminary work by Guo, Zimmerman, and White Brahmia suggests this trend continues in upper division courses [58]. Gains from the 100 to 200 level were reasonably similar across incoming scores, but the 200 to 300 level tells a different story. Students with incoming scores in the top third of the distribution were likely to have higher PIQL scores after taking a 300 level course, whereas students with incoming scores in the bottom third of the distribution were likely to experience little change. A student from the study was reflecting on the score

distribution shown in Figure 6.5, and stated:

> "I found it interesting that even as students advance through higher level physics classes, they don't really gain a better understanding...I honestly found it [the lack of collective improvement] really reassuring...I never felt like I was actually improving, just learning more. It's kind of a debilitation[sic] feeling."

This student appears glad to know that they are not alone, but also describes debilitating feelings. These results suggest that students experiences in a physics major may be partially pre-determined by their incoming PIQL scores, which, like SAT Math scores, are likely to be dependent on the socio-economic status of the school district from which they came [112]. For some, the experience is a negative one.

Research in physics education has traditionally centered on conceptual physics content; research-based interventions tend to focus on a particular conceptual idea that students are likely to find challenging. The development of research-based, instructional interventions that target PQL at the introductory level continues to be a growing body of work within physics education [47, 50, 162, 93], in part because more work is needed to characterize the ways in which conceptual mathematical ideas are used in introductory physics. Recent research in mathematics education has shown that instructional activities targeting conceptual, mathematical reasoning can be effective and helpful towards developing student conceptualization of calculus principles [91, 66]. However, due to the differences between reasoning in mathematics and physics, there is a need to develop similar such materials particular to physics courses.

An essential part of PQL is developing, making sense of, and validating functional relationships between two quantities. Mathematics education research has defined *covariational reasoning* to mean reasoning about how changes in one quantity affect changes in other, related quantities. Covariational reasoning is ubiquitous in physics, describing the reasoning behind statements such as "the force goes like $1/r^2$," "If the speed doubles, the kinetic energy increases by a factor of four," and "as distance from the source approaches infinity,

the potential approaches zero." The language of covariation, as defined in mathematics education research, has been productive towards characterizing student reasoning in physics [137, 46, 104, 90, 6]. There remains a need, however, to characterize what we mean by *physics* covariational reasoning towards developing materials for physics classrooms that help students develop this facet of PQL.

This thesis represents a collection of efforts that aim to better understand how experts and students are reasoning covariationally. The ultimate goal of this work is to provide the necessary research for instructors and researchers to develop instructional materials that support student development of covariational reasoning and PQL more broadly, regardless of their prior instruction. In order to develop such materials, we first need to:

1. Understand what we mean by "covariational reasoning" in physics contexts, as most of the current research has been conducted in mathematics contexts.

2. Identify learning objectives of introductory physics towards covariational reasoning that are within students' zone of proximal development[1].

3. Identify what productive resources students coming into an introductory physics course already have towards reasoning quantitatively in physics contexts.

4. Have an assessment tool to assess potential instruction that builds from where students are towards our learning objectives for a wide array of physics classrooms.

The chapters of this thesis are aligned with these aims. In Chapter 3, I will present a study of covariational reasoning in physics experts. This study was integral to developing an empirical characterization of how physics experts use covariational reasoning differently than mathematicians. Chapter 4 supplements this research by investigating the role that

[1]The Zone of Proximal Development is defined by Vygotsky as "The distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peer" [156].

the meaning of the quantities plays when experts reason covariationally. In Chapter 5, I will present a framework of covariational reasoning that was developed in close collaboration with Alexis Olsho and Suzanne White Brahmia. This framework is emergent from our research, and is designed to be a first step towards helping instructors and other researchers identify learning objectives for covariational reasoning that are appropriate for introductory physics courses.

Chapter 6 focus on two assessment instruments: the Physics Inventory for Quantitative Literacy (PIQL, pronounced "pickle") and the Generalized Equation-based Reasoning Inventory for Quantitative Literacy (GERQN, pronounced "gherkin"). The PIQL was developed several years ago and has been used at the University of Washington since to help assess how students in the calculus-based sequence and beyond are reasoning quantitatively. The GERQN is a newly developed version of the PIQL designed for algebra-based physics classrooms. Chapter 6 describes how PIQL results can be used to assess what resources students have coming into introductory physics courses, and identifies a few resources about covariational reasoning that instructors can build upon. Chapter 6 also discusses the development of the GERQN and preliminary findings from its first year implemented at the University of Washington. It also describes the early findings from an on-going collaboration with KU Leuven, a university in Belgium, as part of our study into comparing how students' quantitative literacy develops during a physics major in Europe and in the U.S.

A version of each research chapter in this thesis has been submitted for publication[2]. I served as the lead researcher and author for all chapters except Chapter 5, however I consider all of this work to be highly collaborative. For Chapters 3, 4, 6, I contributed the majority of the writing. For Chapter 5, my writing contributions are mainly in the background section and in the description of the framework, as well as via regular discussion with the lead author, Alexis Olsho. Alexis Olsho, Suzanne White Brahmia, and I developed the framework together over several years of discussion and consider it a shared result.

---

[2]As of June 2nd, Chapters 3, 6 and **??** have been submitted. We anticipate submitting Chapters 4 and 5 imminently.

The background sections for many of the chapters share similar literary reviews; therefore I have written a brief background section (Chapter 2) to synthesize these ideas for the reader of this thesis. The background sections of individual research chapters have been left in their original forms as a reminder of the important ideas for the studies in that chapter. Finally, I provide a summative chapter (Ch. 7) which provides a brief description of how this work might collectively be used towards developing curricular interventions.

Chapter 2

LITERARY REVIEW

In this section, I aim to situate the work presented in this thesis in the current literature. As this work is focused on reasoning mathematically in physics contexts, I will first discuss research into the differences between reasoning mathematically in mathematics contexts and doing so in physics contexts. I will then share research into how making sense of physical quantities is central to this difference between the two fields.

A core thread of this thesis is in characterizing and assessing covariational reasoning. Therefore, the second section of the background describes current research in covariational reasoning as defined in mathematics education research, and the ways in which physics education research has begun to engage with the language of covariation. The final section provides a brief description of the PIQL, its development, and recent findings upon which much of our analysis and research perspective is based.

2.1 Mathematical Reasoning in Math Class is not the same as Mathematical Reasoning in Physics Class

Research has demonstrated that reasoning mathematically in physics contexts is different from reasoning mathematically in mathematics contexts [130, 13, 84, 67, 116, 163]. In this section, I will provide some examples of this difference, and propose the meaning of quantity as an essential contribution to why students and experts reason differently in physics than in pure math contexts.

In 2016, Bajracharya and Thompson report on students solving a task which asks them to find the change in temperature between two time instances using a graph that depicts the rate of temperature change (dT/dt) on the vertical axis and time on the horizontal axis (Fig. 2.1).

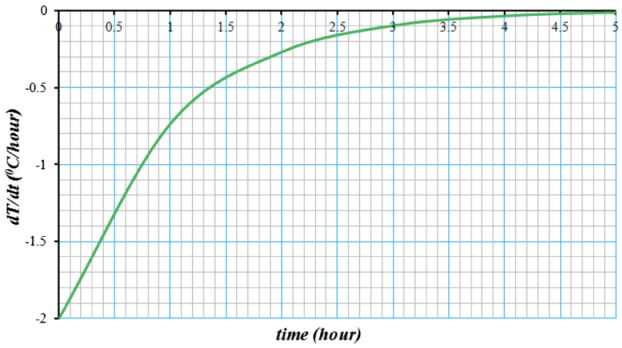


Figure 2.1: The graph used in Bajarcharya and Thompson's study of student quantitative reasoning about "area-under-the-curve" problems [11].

They found that students who did not have a strong understanding of the meaning of the area under the curve could not complete the task. Those that relied on a direct integration approach were stymied by not being provided with a function that related temperature and time, suggesting they have a procedural understanding of integration. Students that had a conceptual understanding of integration were able to connect the procedural and conceptual aspects of solving an integral by finding the area under the curve. This combination of both procedural and conceptual understandings is referred to by mathematics education researchers as *proceputal* understanding, and has been cited as a requisite part of making sense in calculus and physics contexts [56, 17, 104].

Recent work shows that the importance of a proceptual understanding of change persists beyond the introductory level—students that were asked to solve negative integrals wrestled with the meaning of dx representing a small, negative change [10]. In particular, the researchers found that the students had dual meanings of the symbol dx. When asked to make sense of a negative dx quantity, a junior mathematics major stated that it didn't make sense unless viewed in a physical context:

> "When you think about just like, the pure math problems, that's all you really think about—just the fact that dx is just telling you...what variable to use. But...here [in the context of a compressed spring], it represents, it represents something."

This finding highlights the difference between reasoning mathematically in pure math contexts and in physics contexts—physics requires a blended approach, one that combines both the meaning of the quantities involved and the meaning of the procedures that one undertakes to solve a problem.

Redish discusses this difference as a change in dialect; while both disciplines appear to speak the same language, the symbols imply different meaning [116]. In a recent paper, Taylor and Loverude reflect on student views of physics and mathematics. One student is quoted stating:

> "I have math and physics on different days, so I forget about math when I go to physics, I forget about physics when I go to math [136]."

Researchers have argued that at the heart of this difference is the blended nature of physics—in particular, that the physical meaning of the quantities is inseparable from the ways that mathematical reasoning is used in physics [163, 160].

Conceptual blending originated in linguistics to describe the subconcious blending of language and metaphor [52]. It has gained popularity in physics education research to describe the ways that mathematics and physics are inextricably linked [13, 67, 68, 137, 44, 150, 164, 98]. A conceptual blend analysis typically involves a fine-grained approach where mathematics and physics "input spaces" are identified for particular utterances or equations that are subconsciously blended into one idea. There is also growing use of the term more informally amongst physics education research to refer to the blended nature of mathematical structure and physical meaning [166, 103, 44, 86]. The work presented in this thesis takes the perspective that mathematical reasoning in physics is inherently a blended cognitive activity.

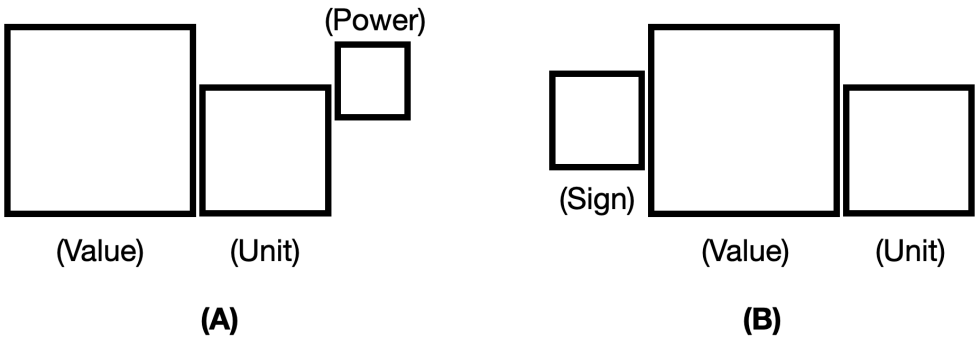


Figure 2.2: (A) the measurement symbolic form [42], (B) the quantity symbolic form [160]

2.2 Reasoning about Symbols and Quantities in Physics Contexts

Research in both physics and mathematics education has shown that the meaning of quantity plays an essential role in developing, making sense of, and validating mathematical models [139, 71, 163, 165, 39, 40, 78]. However, armed with the knowledge that mathematical reasoning in physics is distinct, there is a need to better understand how physics students and experts do use and make sense of symbols and quantities.

Sherin developed a framework to characterize symbolic patterns in the ways that upper-division physics students were observed to make sense of various physics equations [130]. These "symbolic forms" represent how a proceptual, and blended, view of symbols and quantities are useful in physics education. For example, consider the symbolic form *base + change*: $\square + \Delta$. In this form, Sherin identified how students described a quantity accumulating, beginning with an initial value and gaining amounts of change over time. Mathematics education research, in which quantity often means a value with a unit, has defined a *measurement* symbolic form shown in Fig. 2.2A. White Brahmia adds to this definition in physics to develop the *quantity* symbolic form which includes sign, and implies a physical connection to the meaning of the quantity that goes beyond the unit associated with the value (Fig. 2.2B) [160]. For example, consider a variable m which is understood to have units of kilograms. In physics, the unit kilograms also implies that the value of m is understood to be a positive

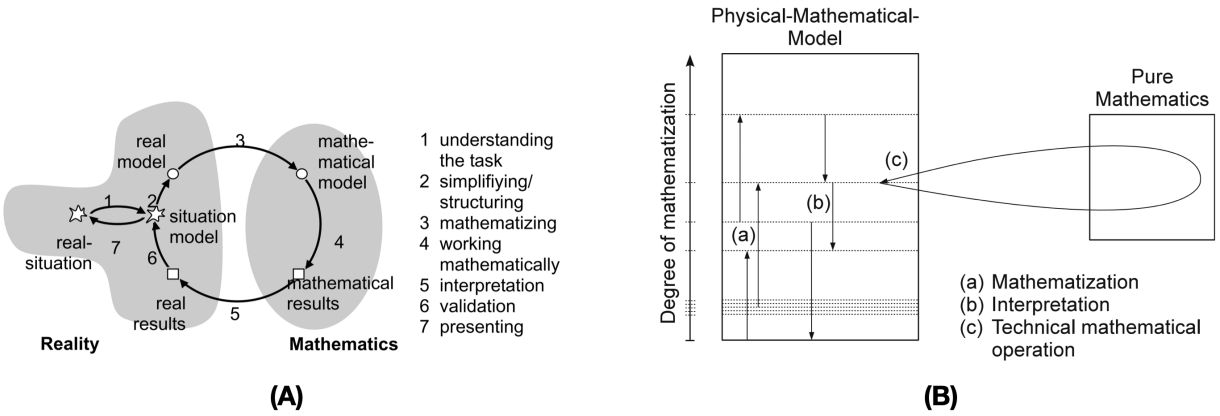


Figure 2.3: Two popular modeling cycles in mathematics and physics education research; (A) by Blum and Leiß[16] and (B) by Uhden, Karam, Pietrocola, and Pospiech [148]

number.

Reasoning about symbols in the context of physics has been shown to be more difficult for students coming into physics from mathematics courses, where the physical meaning of the quantities cannot be guaranteed to "come along for the ride" [130, 143, 163, 103]. Traditionally, the development of mathematical models is referred to as a cycle, where the first and last steps are often centered on connecting the mathematics to the meaning of the quantities involved. Consider the two modeling cycles in Figure 2.3 that are popular in both mathematics and physics education. Common across these different descriptions are two important features: (1) the separation of physics and math reasoning spaces, and (2) the cycles that are shown to leave from, and return to, the physical world between which reasoners "do math." However, making sense of the meaning of the quantities has been shown to be an essential, continuous activity when modeling the real world [37, 160].

Research in mathematics education suggests that there is little evidence students reason in cyclic patterns [53, 7]. Recent work by Czocher examined engineering students solving a wide variety of word problems akin to those in mathematics and engineering textbooks. Her research showed that successful students—those that were able to produce mathematical models of the various problems they were asked to solve—did not follow a cyclic approach

as suggested by these diagrams [37]. Rather, the students were shown to go back to the problem to reflect on the meaning of the quantities involved throughout model development (rather than simply at the end). Similarly, Serbin and Wawro demonstrate that upper-division students enrolled in quantum mechanics physics courses were likely to use "physics-like" and "math-like" reasoning in deeply interconnected ways, as opposed to reasoning first physically, then computing, and then making sense of the mathematics in the context of the problem [128]. The authors state "when students draw on their mathematical knowledge... it relies on their understanding of how the mathematical concepts and physics concepts are intertwined" [128]. We expand on our own observations of the connection between quantity and mathematical sense making in Chapters 4 and 5.

2.3 Covariational Reasoning in Mathematics and Physics Education

Covariational reasoning—reasoning about how one quantity changes as a result of changes in another, related quantity—is an essential part of reasoning quantitatively in physics and developing, making sense of, and validating mathematical models [28, 141, 166]. Covariational reasoning was developed in mathematics education research, and is an essential tool for reasoning about a relationship between two or more quantities [138, 36, 125, 28, 94, 33, 70, 65, 141, 107, 45, 23]. In recent years, physics education research has taken up the language of covariation [137, 46, 104, 90, 6]. There is also a large body of work in physics education that relates to covariational reasoning prior to the adaptation of the language from mathematics education research [18, 5, 88, 130, 8, 146, 15]. In this section, we provide a brief summary of how covariational reasoning has been used in mathematics education research and in physics education research.

2.3.1 Covariational Reasoning in Mathematics Education Research

Covariation has been studied in contexts of reasoning about function [36, 29, 100, 72, 141, 107], as well as graphing and the use of coordinate systems [95, 157, 70, 65, 24], and found to be an essential component of reasoning about rates of change. Mathematics education

researchers Carlson et al. developed a framework of *mental actions* (MA) and associated behaviors that effectively operationalizes covariational reasoning [28]. This original framework was based on results from a study in 1998 into undergraduate and graduate students reasoning about function [28, 29]. The framework was revised in 2017 [141], and summarized by Jones in 2022 [76]. These frameworks have been productive for studies into how students reason about rates of change [28, 31, 71, 76], limits and accumulation [30], coordinate systems [95], parametric relationships [106], and generating graphs [33, 157, 65]. We find Jones' simplified and streamlined framework was most productive for our analysis across the studies presented here [76]. A slightly modified version is shown in Table 2.1.

2.3.2 Covariational Reasoning in Physics Education

In physics education research, reasoning about how two or more quantities change with respect to one another often falls under the names of proportional reasoning or scaling [5, 18, 17, 15]. Proportional reasoning typically refers to directly proportional relationships (i.e. $F \propto a$), and has at times been extended to refer to non-linear relationships (i.e. $U \propto -1/r$). Scaling is often considered to refer to geometric contexts, however it is also stretched throughout the literature to refer to relating discrete changes of two quantities (i.e. if I double this, what happens to that?). In the language of covariational reasoning, we consider proportional reasoning to be a *linear version* of covariation and scaling to be an instance of *discrete* covariation. Recent work in physics education research has demonstrated that the language of covariational reasoning from mathematics education is helpful in analyzing novice and expert work [137, 46, 104, 90, 151, 172, 171, 133]. However, preliminary work in characterizing physics experts' covariational reasoning through the lens of covariation as defined in Table **??** suggests that there are some distinct differences between the ways that physics and mathematics experts reason covariationally [172, 171].

Label	Mental Action [28, 141]	Brief Description [76]	Example Behavior
MA 1	Recognize Dependence	Identify variables that are dependent	Labeling axes
MA 1.5	Precoordination	Asynchronous changes in variables	Articulating that first, one quantity changes, and then the other changes
MA 2	Gross Coordination	General increase/decrease relationship	Describing that as one quantity increases, another decreases
MA 3	Coordination of Values	Tracking variable's values	Plotting points
MA 4	Chunky Continuous	Values changing in discrete chunks	Articulating that as one quantity doubles, the other triples
MA 5	Smooth Continuous	Continuous, simultaneous changes	Describing that the quantities vary together, smoothly and continuously

Table 2.1: A summary of the covariational reasoning mental actions (MA) used in this study. The framework we provide here is a slightly modified version of that summarized by Jones in 2022 [28, 141, 76]. We use names and descriptions from Jones' summary. Numeric labels and example behaviors were added by the authors and correspond to the original mental actions from the 2002 framework, with the addition of 1.5 for "precoordination", for ease of reference within this paper.

You are purchasing a slide for a playground and would like to get the steepest one you can find. For four different slides, you have the measurements of the length of the base of the slide (measured along the ground), and the height of the slide.
You decide to use this information to rank the slides from **most steep** to **least steep.** Which of the following choices is the best ranking?

a. A > B = C > D
b. B > C > A > D
c. A = B > C > D
d. B > A = C > D
e. A = D > C > B

Slide	Base	Height
A	8 ft	12 ft
B	5 ft	9 ft
C	6 ft	9 ft
D	12 ft	8 ft

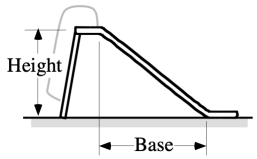


Figure 2.4: An example item from the PIQL [166, 1].

2.4 Assessing Quantitative Literacy: the PIQL

In order to determine whether instruction is improving students' quantitative literacy, including covariational reasoning, we require an assessment tool. The Physics Inventory for Quantitative Literacy (PIQL) was developed at the University of Washington, and is available for widespread use [166, 1]. The PIQL is designed for calculus-based classrooms, and is a *reasoning* inventory, meaning that it measures how students make sense of novel situations rather than content knowledge. An example of a PIQL item is shown in Figure 2.4. In this item, students are prompted to develop a way to quantify the "steepness" of the slide, and rank four slides based on how steep they are.

The items on the PIQL are designed to measure three main categories of quantitative reasoning: reasoning about signs, proportional reasoning, and covariational reasoning. We note that while proportional reasoning is a subset of covariation, it is so common in physics that we consider it its own category for the purposes of item development. The PIQL was developed and validated over several years, following the procedure described by Adams and Weiman [3]. It has been administered in the calculus-based sequence for several years at the University of Washington, and recent work has begun exploring how students' scores change in upper division courses [**?**].

Analysis has found that PIQL scores are unlikely to change during the course of instruc-

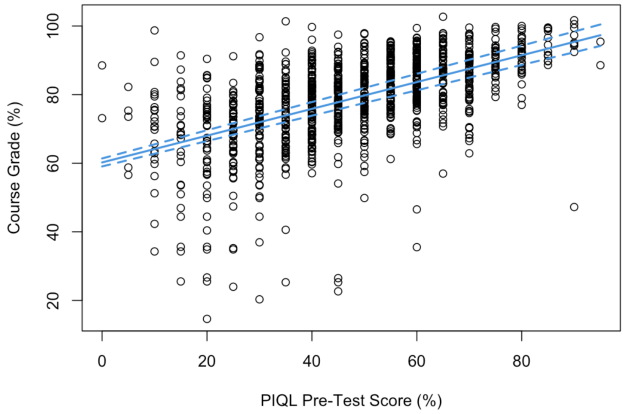


Figure 2.5: The correlation between incoming PIQL score (before any physics instruction at the University of Washington) and final course grade after one quarter of introductory physics. These data represent a sample of 1024 students. The linear regression shown is described by: (Course Grade) $= (60 \pm 1) + (0.39 \pm 0.02)$(PIQL Score). $R^2 = 0.31 \pm 0.03$.

tion (Fig 1.1 A), and are positively correlated to course grade (Fig 2.5). A single regression analysis shows that $R^2 = 0.31 \pm 0.03$, which can be interpreted to mean that 30% of the students' course grade can be explained by their incoming PIQL score. If quantitative reasoning is not improving as a result of instruction—especially considering the instruction at the University of Washington has been developed over many years as a result of research—there is a concern that participation and perceived success in physics may be related to the school district students are coming from.

We suspect this issue continues in the algebra-based sequence, and therefore have developed an algebra-based version of the PIQL. This assessment's development, distinction from the PIQL, and early results are discussed in Chapter 6.

Chapter 3

CHARACTERIZING EXPERT PHYSICS COVARIATIONAL REASONING IN GRAPHING TASKS

3.1 Introduction

Learning to model the physical world quantitatively is a key objective of courses in physics, math, and other mathematics-based disciplines [63, 48, 20]. Reasoning quantitatively is at the heart of what it means to "think like a physicist." There is a large body of research in physics education that has examined the role of conceptualizing mathematics in how students reason about physics [148, 79, 154, 25, 69]. Emergent from this research is evidence that quantitative reasoning is fundamentally different in a physics setting than in a purely mathematical one [130, 13, 84, 67, 116, 163].

A central construct for this chapter is *covariational reasoning*. Covariational reasoning is a common subject of mathematics education research [138, 28, 102, 70, 141, 65, 76]. A simple working definition is that covariational reasoning is the reasoning used when coordinating how two quantities vary. Mathematics education researchers have effectively operationalized covariational reasoning by creating a framework of "mental actions" that describes specific ways in which covariational reasoning can be observed in novice and expert mathematical reasoning [28, 141, 76]. Covariational reasoning, and the associated mental actions described by mathematics education research, resemble ideas widely found in physics; e.g. in the immediate moments after a switch is closed in a circuit, the voltage increases as time increases. However, the rate of increase diminishes until the voltage reaches a steady value equal to the battery voltage.

The work described in this chapter began as an effort to determine the extent to which covariational reasoning, as described by mathematics education researchers, is a useful lens

through which physics reasoning can be studied. We began by replicating a study of expert mathematics covariation with physics experts and comparing the two populations' reasoning. The replicated study involved novel graphing tasks—that is, the tasks prompted participants to relate familiar quantities in unfamiliar ways. The lack of familiarity with the relationship between the quantities was intended to prompt participants to use covariational reasoning. As our project with physics experts progressed, it became clear that covariational reasoning as described by mathematics education research is indeed a useful way to analyze reasoning by expert physicists, but that *physics* covariational reasoning encompasses a broader set of thought processes.

Therefore, we expanded our investigation into two questions:

- **Research Question 1:** In what ways are the covariational reasoning frameworks developed by mathematics education researchers productive for describing expert physicists' modeling in novel graphing tasks?

- **Research Question 2:** What patterns do we observe in the covariational reasoning of physics experts as they solve novel graphing tasks?

We examine these questions across three studies: Study 1, in which we replicate mathematics education researchers Hobson and Moore's 2017 study of expert mathematics covariational reasoning [65] with physics experts; Study 2, an extension of this work with tasks situated in more advanced physics contexts; and Study 3, an extension of Studies 1 and 2 with a broader population. We identify the ways in which the mathematics covariational reasoning framework is useful in analyzing physics experts' reasoning, and propose five emergent reasoning devices and three problem solving modes that represent a wider set of thought processes related to physics covariational reasoning. While these devices and modes may not be exclusive to physics, they were prevalent throughout our experts' reasoning and seem essential to a complete description of physics covariation.

3.2 Background

In this section, we situate our work in the current literature on covariational reasoning in mathematics and physics education research. As covariational reasoning is an essential part of developing, making sense of, and validating mathematical models [139, 28], we begin by reflecting on traditional perspectives on the mathematical modeling in physical contexts and how these perspectives relate to this work. We then discuss covariational reasoning as defined by mathematics education researchers, and examine work in physics education research on reasoning about rates of change through a covariational reasoning lens. Finally, we provide a brief summary of expert mathematics covariational reasoning as described by Hobson and Moore [65], as their study forms the jumping off point of our research.

3.2.1 Modeling and Covariational Reasoning

Modeling involves making connections between the real world and mathematics, as shown in Blum and Leiß's modeling cycle (reproduced in Fig. 2.3A) [16]. Research in mathematics and physics education demonstrates that attending to the physical meaning of a quantity is an essential part of modeling across many contexts [139, 161, 160, 39, 78]. Reasoning about the meaning of quantities in a particular model is traditionally described as an initial and/or final step in the modeling cycle. However, a recent study of engineering students demonstrated that making sense of the meaning of the quantities involved in a particular model was a continuous activity throughout the modeling process [37]. These findings are in line with recent work that suggests reasoning mathematically in physics can be viewed as a blend of conceptual mathematical ideas and reasoning about the physical meaning of the quantities involved [13, 67, 116, 137, 68, 44, 160, 150]. We take the blended nature of physics and math as a foundation of our research perspective. Therefore, we suggest that *physics* covariational reasoning is necessarily grounded in the meaning of the relevant physical quantities.

In mathematics education research, covariational reasoning is defined broadly as reasoning about how two or more quantities vary with respect to one another [141]. It has been

studied in the context of mathematics for several decades, and is an essential tool for reasoning about a relationship between two or more quantities [138, 36, 125, 28, 94, 33, 70, 65, 141, 107, 45, 23]. Mathematics education researchers Carlson et al. developed a framework of *mental actions* (MA) and associated behaviors that effectively operationalizes covariational reasoning, and has been used by researchers to investigate covariational reasoning in student and expert statements [28, 141, 65, 76]. This original framework was based on results from a study in 1998 into undergraduate and graduate students reasoning about function [28, 29]. The framework was revised in 2017 [141], and summarized by Jones in 2022 [76]. We find Jones' simplified and streamlined framework was most productive for our analysis [76], and a slightly modified version is shown in Table 2.1.

Physics education researchers describe student reasoning about change using the language of proportional reasoning [18, 5, 88, 130] and scaling [8, 146, 15]. In the context of covariational reasoning, we argue that proportional reasoning is a *linear version* of covariation and that scaling could be described in the language of the mathematics covariational reasoning framework as Mental Action 4: Chunky Continuous, in which one might ask "if I double this, what happens to that?" Physics education research has already begun to benefit from integrating the language of covariational reasoning in investigations into mathematical modeling in physics [137, 46, 104, 90, 6]. Incorporating the language of covariation has thus far been useful in physics education research, and we suggest this may be because it allows for descriptions of reasoning about change across a wide variety of continuous, functional relationships between two or more quantities [76].

Covariation in mathematics is centered on explicitly coordinating the change in one quantity with the change in another, i.e. "as time increases, the distance will also increase." We will refer to this as "direct" covariational reasoning. While direct covariational reasoning certainly appears often in physics, there is some preliminary evidence that covariation and reasoning about quantities more generally may have additional, distinct features in physics [172, 160]. In particular, early findings in our prior work suggest that physics covariational reasoning may involve a broader tool set than reasoning focused on the specific functional

relationship between two quantities [172, 104], perhaps due to the blended nature of physics reasoning. In this chapter, we will use the mental actions as described in Table 2.1 as a launching point for our investigation into physics covariational reasoning.

3.2.2 Expert Mathematics Covariational Reasoning: Hobson and Moore's 2017 Study

The first study described in this chapter is a replication of a study conducted by Hobson and Moore in 2017 [65]. That study examined how expert mathematicians use covariational reasoning by asking several mathematics graduate students to complete novel graphing tasks. The authors considered graduate students to represent experts in the context of the introductory calculus problems, as graduate students are expected to "have sufficient experience engaging in tasks involving representing and reasoning about quantities" [65]. The graduate students were given animations developed by the researchers of three graphing tasks identical to those shown in Figure 3.1, with place names adjusted to the local region.

All three tasks asked participants to relate a particular quantity (e.g., the height of a Ferris wheel cart) to the total distance traveled of an object moving at constant speed. This was an intentional choice by the researchers, who noted the facility with time-based motion tasks that many experts in math-based disciplines have acquired. The research demonstrated that mathematics graduate students were likely to:

(1) Directly covary the quantities asked for in the task (in some cases, with prompting from the interviewer), and

(2) Equally divide the domain of the task, or large sections of the motion, into small sections to compare the change in the two quantities and then map these equal segments to their graphs [65].

In this chapter, we report the results from our replication study and compare our observations to those provided by Hobson and Moore. We then discuss the results from the extension studies that examined a broader set of physics contexts and included a larger population.

	Study 1: Replication	Study 2: Extension	Study 3: Faculty
Subjects	10 Graduate Students	5 Graduate Students from Study 1 5 New Graduate Students	5 R1 Faculty
Tasks	Going Around Tacoma	Gravitation Electric Charge	Gravitation
	Ferris Wheel Square Track		Ferris Wheel
		Drone	Drone
		Intensity	Intensity

Table 3.1: A summary of the differences between the populations and tasks across the three studies.

We report on the patterns we observed across all the studies as physics experts reasoned covariationally.

3.3 Methods

In this section, we describe the methods, population, and interview tasks used to conduct the series of studies. We discuss the two analytical approaches we used in our coding process—grounded theory and thematic analysis—to account for the difference between Study 1, during which we were informed only by our literary review, and Studies 2 and 3, during which our analysis was influenced by our experiences from Study 1. Finally, we describe how we sought patterns in the coded data to address our two research questions.

3.3.1 Study Design

This work was situated at a large, R1 university in the Pacific Northwest. The population across the three studies included 20 experts, nine of whom were female or non-binary identifying and 11 of whom self-identified as male. The pool of experts spanned from first-year graduate students through tenured professors with over 30 years of experience. The

population also spanned a variety of physics subfields, including both experimentalists and theorists. An effort was made to include diversity across race and country of origin. We categorize the graduate students into two groups: those who were early in their graduate career and were still enrolled in required courses, and those late in their graduate career who were fully engaged in research. We entered the research study assuming all of our participants to be experts in the physics contexts of the tasks across all three studies. We discuss how the populations compare in the findings.

Study 1 was a replication of the Hobson and Moore 2017 investigation into expert covariational reasoning that was conducted with mathematics graduate students (see Sec. 3.2.2). In order to study expert physics covariational reasoning, we repeated their study with a population of physics graduate students. Study 2 was an extension of Study 1, using new tasks situated outside of constant motion contexts. Constant motion contexts are so commonplace for physics graduate students that they are likely part of their everyday thinking. We hoped to elicit a wider variety of reasoning by asking the graduate students to engage with less-familiar contexts in Study 2.

Study 3 was conducted with five tenure-track or tenured physics professors. The intent was to see if the patterns we noticed in graduate students' reasoning were present in faculty members' reasoning, and if any new patterns might emerge. Study 3 used a subset of tasks from the Studies 1 and 2 (see Table 3.1).

In all three studies, participants were solicited by email and offered a gift card to a local coffee shop as a thank you for their time and effort. An interview involved one interviewer and one participant. During each interview, the participant was presented with a handout of three to four graphing tasks. Participants had access to a computer that allowed them to interact with short animations associated with each prompt. Each task prompted the participant to create a graph to represent the relationship between two quantities. The participants were invited to move between questions at their own pace, and play the animations as desired. The interviews lasted between 25 and 50 minutes, depending on how long the participants chose to work on the tasks. The participants were encouraged to think aloud while they created their

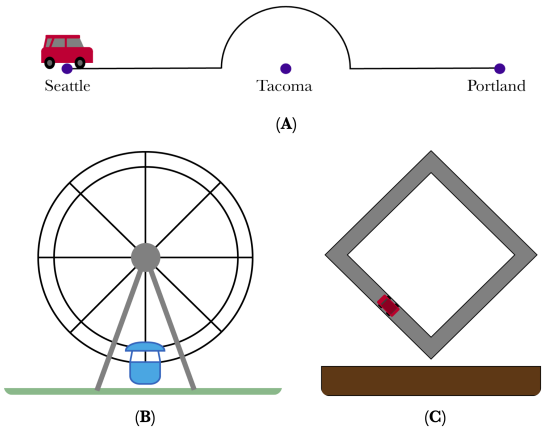


Figure 3.1: Stills from the animations associated with tasks administered in Study 1: (A) Going Around Tacoma, (B) Ferris Wheel, and (C) Square Track. All tasks asked participants to create a graph that related a particular quantity (distance to Tacoma, height of the cart, and distance from the cart to the wall, respectively) to the total distance traveled of an object in constant motion. These graphing tasks are identical those in the original study conducted by Hobson and Moore [65], with the exception that we changed the place names in Going Around Tacoma.

graph, and the interviewer asked follow up questions after the participant expressed they were finished with all of the tasks. The interviews were audio recorded; initial transcripts were created automatically using the Otter.ai software program [2], and were subsequently hand corrected.

The following sections discuss details particular to each study, including a brief description of the tasks. A summary comparison of the populations and tasks used in the three studies can be found in Table 3.1.

3.3.2 Study 1: Replication Study

In Study 1, we followed the procedure described in Hobson and Moore's 2017 paper. We used the same tasks as those they administered with minor modifications of the place names to reflect the local area (Fig. 3.1). We included both written descriptions and animations of each task's context. We recruited 10 physics graduate students to participate as our experts,

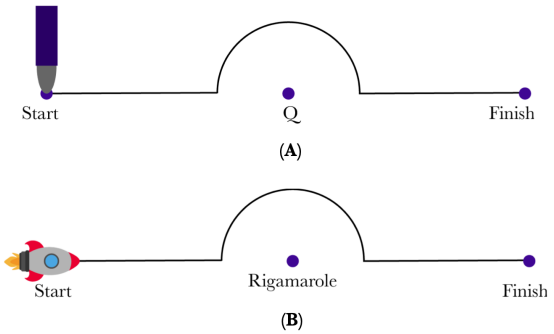


Figure 3.2: Stills from the animations associated with the first task administered in Study 2. Participants either saw (A) Electric Charge or (B) Gravitation. The tasks prompt participants to create a graph that relates either the electric or gravitational potential and the total distance traveled of the probe or spaceship, as it moves at constant speed from start to finish.

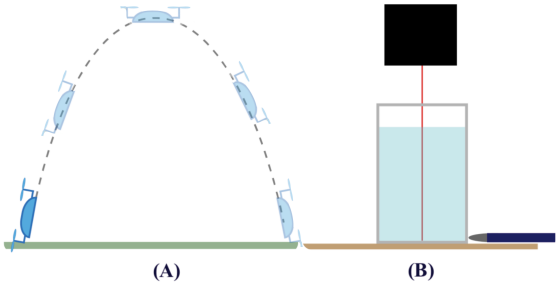


Figure 3.3: Stills from the animations associated with the second and third tasks administered in Study 2: (A) Drone, in which participants were prompted to create a graph that relates the height to the angle of the drone, flying in the arc shown, and (B) Intensity, in which participants were prompted to create a graph that relates the intensity measured by the probe at the side of the liquid column to the total distance traveled of the probe as it moves from the bottom of the column to the top.

consistent with the original study [65].

Each task prompts the participant to create a graph that relates a particular quantity to the total distance traveled by an object (Fig. 3.1). In the first task, Going Around Tacoma, participants are asked to relate the distance from Tacoma and the total distance traveled by the car. In the second task, Ferris Wheel, participants are asked to relate the height of the Ferris wheel cart and the total distance traveled by the cart. In the third task, Square Track, participants are asked to relate the distance of the go-cart from the wall to the total distance traveled by the go-cart.

3.3.3 Study 2: Extension Study

The extension study was designed upon reflection of the results from the replication study (Study 1). The interviewers noticed that the physics graduate students were so well acquainted with motion-based tasks that few of the graduate students engaged in direct covariational reasoning—that is, they were unlikely to compare the change in one quantity to the change in another, and instead assigned functional relationships between the two quantities [172]. While it may not surprise physics instructors and researchers that an expert would rely on their understanding of kinematics rather than develop a new relationship between the two distances, this was notably different from the mathematicians. The similarities and differences in the patterns from our interviews are compared to those reported by Hobson and Moore thoroughly in Section 3.4.1.

In order to better understand how physics experts use covariational reasoning, a new set of three tasks were developed using physics contexts outside of kinematics. A summary of the tasks can be found in Figures 3.2 and 3.3. We designed two isomorphic versions of the first task, Going Around Tacoma. One version, Electric Charge, asks the participants to relate the electric potential measured by a probe and the total distance traveled as the probe moves around a charge (Fig. 3.2A). The other version, Gravitation, asks participants to relate the gravitational potential of the system and the total distance traveled of a rocket ship as it moves around a planet called Rigamarole (Fig. 3.2B). Half of the graduate students

were given Electric Charge, and the other half were given Gravitation.

The second task, Drone, was designed to relate two familiar quantities—height and angle—in a novel way (Fig. 3.3A). The task asks participants to relate the height of a drone to the angle the drone makes with the horizontal axis, with the constraint that the drone is designed to tilt such that its axis is always aligned with its velocity vector (as shown with the faded drone images in the figure). Finally, the third task, Intensity, depicts a laser shining through a column of liquid and asks participants to relate the intensity measured by a probe and the total distance traveled by the probe as it moves with apparent constant speed upward along the column of liquid (Fig. 3.3B). All of the graduate students saw both Drone and Intensity.

Five graduate students from the first study elected to participate in the extension study, and five additional graduate students were recruited to ensure there was no bias from participating in the first round.

3.3.4 Study 3: Faculty Study

The final study probed whether the patterns that emerged amongst graduate students could be observed in faculty reasoning and whether any new patterns emerged in a population with more extensive physics experience. The tasks for this study were chosen based on the researchers' experience with the tasks in the first two studies. Ferris Wheel, Gravitation, Drone, and Intensity had been most successful at eliciting covariational reasoning.

Five faculty members were recruited individually by email based on their strong engagement with teaching. We considered engagement with teaching as a measure of those faculty that may be most likely to be able and willing to think aloud about their reasoning.

3.3.5 Data Analysis

The interviews were analyzed through an iterative, qualitative coding process. For Study 1, the research team's analysis was informed by prior literature on covariational reasoning in mathematics education research, as well as prior literature on proportional reasoning and

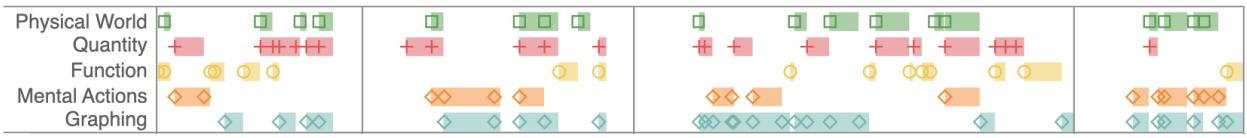


Figure 3.4: An example time on task chart for a faculty member from Study 3. The horizontal axis represents normalized time, or the fraction of time on task. The symbol represents the time the category was coded, and the band represents the duration that the interviewee was engaged with that kind of reasoning.

scaling in physics education research. The aim of Study 1 was to investigate the similarities and differences in the ways the physics graduate students reasoned covariationally, as compared to the covariational reasoning of mathematics graduate students reported in Hobson and Moore's study. However, we wanted to ensure that the results rose out of the data itself. We chose to set aside the mathematics covariational reasoning framework and employed a modified grounded theory approach [55] to seek patterns in the data. A grounded theory approach provided a framework to acknowledge the prior knowledge the research team brought to the analysis and to reflect the effort made to seek emerging codes from the data.

Grounded theory involves three main steps: initial coding, intermediate coding, and advanced coding [55, 105]. In initial coding, the research team codes as they read through the data, allowing the codes to emerge from the data as they read it. This process is refined and reviewed through iterative passes of coding until the codes are stable and orthogonal—meaning every code is applied more than once, and that each code means something different and distinct than another. These codes are then reflected on by the research team, and reliability is ensured by having a second member of the research team code the data using the proposed codebook. If the codes are consistent, the codebook is considered validated. During intermediate coding, the codes are collected into categories. During our analysis, this process was conducted by having two members of the research team group codes separately, and then come together to compare, discuss, and finalize the code categories. Finally, during advanced coding, the entire research team reflects together on the categories to find patterns

and a cohesive story throughout the data.

Studies 2 and 3 were conducted after Study 1 was analyzed. A thematic analysis framework [97] was used to account for the ways in which the results of Study 1 impacted the research teams' interpretation of the data. Thematic analysis provides a structure for the research team to seek emergent patterns while keeping a record of how prior experiences with the subject may be influencing the choices they make.

Thematic analysis involves similar steps to grounded theory, but the process is expanded to include additional reflection by the entire research team and to address issues of prior experience [97]. The first step of thematic analysis, for example, is to gain familiarity with the data. In contrast to grounded theory, where one codes during an initial first pass of the data, thematic analysis encourages the entire research team to become familiar with the data. During this process, documentation of thoughts and reflections is essential to maintain a written record of research. During the initial coding phase, thematic analysis relies on peer debriefing—continuous discussion across research team members—and researcher triangulation, during which multiple members of the research team reflect on pieces of transcript to choose codes. Similarly to grounded theory, this process is an iterative one. However, discussion amongst the research team is encouraged to ensure validity of the codes. This emphasis on discussion and documentation continues through the categorization and pattern recognition phases.

To ensure the codes were grounded in the data, research team members met regularly during the coding process to discuss how the data supported the particular codes that emerged. Studies 2 and 3 were coded together using the same codebook.

For each set of studies, the lead researcher cyclically coded the transcripts for emergent patterns related to quantitative reasoning. Between coding cycles, the lead researcher conferred with the research team to establish inter-rater reliability and wide-spread agreement with the coding scheme. We established code categories by independently clustering codes, and then discussing our clusters as a group until consensus was reached. The resulting code categories that are pertinent to the discussion in this chapter can be found in the Appendix.

The codes and code categories allowed us to compare small pieces of reasoning to the mathematics covariational reasoning framework, and begin to pull apart the similarities and differences between the reasoning of mathematics and physics experts. For example, when we compared the emergent code categories from Study 1 to the framework in Table 2.1, we found some important connections between the codes we identified and the mental actions in the mathematics covariational reasoning framework (see Section 3.4.1). As a result, we used these mental actions as codes in Studies 2 and 3.

It was also evident that there were larger patterns across the experts and tasks for all three studies. Therefore, we sought patterns between code categories to characterize the ways in which physics experts engaged with covariational reasoning as part of the wider group of strategies they used to solve the tasks. We approached this pattern analysis differently for Study 1 and for Studies 2 and 3 to account for our prior experience; our process for each is described in detail in the following sections.

Study 1 Data Analysis

To seek patterns between code categories for Study 1, two members of the research team performed a card-sorting task [34, 127] to characterize larger patterns in expert reasoning. They separately grouped sections of coded transcript based on the ways participants expressed their reasoning. The two researchers compared their findings and compiled a list of reasoning patterns that were observed across all participants. Each pattern is associated with one or more code categories.

Studies 2 and 3 Data Analysis

Studies 2 and 3 were led by the same researcher who led Study 1. In order to mitigate confirmation bias while interpreting the patterns in our code categories, we opted out of using a card sorting task and instead created a timeline (see Figure 3.4) for each participant. The timelines are plots that show the code categories assigned to a participant over time. Each plot shows the fraction of total time on task along the horizontal axis, and a mark for

each code category at the each moment it was assigned. The tails associated with each mark show for how long the participant was engaged in that code category. Each code category is plotted with a different shaped mark and color. Therefore, this method allowed the research team to find patterns between colors in the diagrams and then draw conclusions about how code categories may be interacting, grounding the analysis in the data.

The research team sought patterns in the code categories by looking for categories that occur at the same time and categories that rarely appear together. As an example, consider the patterns in Figure 3.4. The code category "Quantity," which includes symbolizing and relating quantities to one another, often appears at the same time as the code category "Physical World," which includes making sense of the physical context of the task. In contrast, the code category "Function," which includes assigning functional forms, often appears next to—but not at the same time as—the code category "Mental Actions," which includes the mental actions from the mathematics covariational reasoning framework.

Once initial patterns were identified across the 15 interviews in Studies 2 and 3, we went back into the transcripts to examine how the participants described their thought processes and what behaviors were observed by the interviewer at those moments where code categories coincided or were apart. We then iteratively went back and forth between the timelines and the transcripts to characterize the patterns we observed.

During this time, some new patterns emerged, some were revised and clarified, and others were set aside. For example, the code category "Graphing," which describes behaviors associated with drawing a graph, often appears at the same time as "Mental Actions" in Figure 3.4. The transcript reveals, however, that this is because the mathematical mental actions describe reasoning that is closely tied to graphical behaviors (e.g. identifying that two quantities are related is tightly associated with the behavior of labeling the axes of a graph). This relationship is well documented in mathematics education research, and taken as a premise of the mathematical mental actions. Therefore, we did not further characterize this relationship.

3.4 Findings

Three main patterns emerged:

A. Physics experts engaged with covariational reasoning in some ways that align with the mathematics covariational reasoning framework (towards Research Question 1).

B. Physics experts used a set of five distinct "reasoning devices" that were not well described by the mathematics covariational reasoning framework (towards Research Question 2).

C. Physics experts broadly engaged in three different modeling modes when they were solving covariational graphing tasks (towards Research Question 2).

During our analysis, we also found empirical evidence of the intertwined nature of physical sense-making and mathematical modeling. This integrated way of reasoning about math and physics was an essential part of how all experts engaged with the tasks. In this section, we first discuss the relationship between math and physics, as it underlies the rest of our analysis. We will then discuss the main findings in the subsections that follow.

While we did not intend to investigate the relationship between the meaning of the physical quantities and conceptual mathematics, it was a consistent finding that was integral to how the experts reasoned covariationally. As the example timeline in Figure 3.4 shows, the "Physical World" and "Quantity" code categories often appeared at the same time. Here, our code category "Physical World" represented when the experts were relating their reasoning back to the context of the task. "Quantity" referred to when experts symbolized a quantity, discussed its mathematical structure, or related one quantity to another. In particular, we found that 78% of the time that Quantity was coded, Physical World was as well. While this makes sense because they are reasoning about covariation in a physics context, we note that mathematicians rarely do this.

The majority of the time these codes coincided, it was because the research team could not separate the two ideas within a single utterance. For example, one expert working on Gravitation stated:

> "It's got to be negative one over R. Because I want it to fall in."

We stated in Sec. 4.2 that we take the blended nature of physics and mathematics as a foundation of our research perspective and our evidence supports this perspective. We suggest that the rest of the findings presented here should be interpreted as all involving a blend between physics and mathematical reasoning. This finding warrants a more thorough discussion than is appropriate here, and we are preparing a subsequent report to describe our findings in detail.

3.4.1 Seeking Mathematics Mental Actions in Physics Expert Reasoning

Research Question 1 states, "In what ways is the covariational reasoning framework, as defined by mathematics education researchers, productive for describing how expert physicists reasoned while solving novel graphing tasks?" We found that the mathematics covariational reasoning framework (Table 2.1) was helpful to describe the physics experts' reasoning in all three studies.

In Study 1, we made an effort to set aside the mathematics covariational reasoning framework to seek emergent codes from the data. Once the data were fully coded, we were able to draw connections between the codes and some of the mental actions described in the mathematics covariational reasoning framework. We found that MA 1 (Relating Quantity), MA 2 (Gross Coordination), MA 3 (Coordination of Values), and MA 4 (Chunky Continuous) were all present in the codebook (see the Appendix for more details about the codes themselves). However, we note that we were unable to relate any of the codes that emerged from our analysis to MA 5 (Smooth Continuous).

In Studies 2 and 3, we used MA 1-4 as codes under the code category "Mental Actions." These codes were very productive towards characterizing physics experts' reasoning.

However, there were other code categories which we consider to be related to the broader definition of covariational reasoning—how the changes in one quantity are related to changes in another quantity—but are not well described by the mathematics covariational reasoning framework.

Research Question 2 states, "What patterns do we observe in the reasoning of physics experts as they solve novel graphing tasks?" We identified five "reasoning devices"—thought processes and behaviors that experts used in while reasoning covariationally—and three larger modes that physics experts engaged in to create graphical models. These reasoning devices and modes are distinct from MA 1-5, but are all representative of covariational reasoning more generally. The following two sections define the devices and modes, and discuss how they are related.

3.4.2 Reasoning Devices

During the card-sorting task for Study 1, we identified five patterns of expert reasoning that are only partially described by the mathematics covariational reasoning framework. These patterns were consistent across populations and tasks during Studies 2 and 3, and they encompass many of the code categories that were different than the mental actions. We call these "reasoning devices," to illustrate that they are a combination of how the experts described their own reasoning while talking out loud and the graphing behaviors we observed the participants perform. In the following subsections, we describe each of these reasoning devices and the associated mathematics mental action (MA 1-4), if any. A summary of the reasoning devices can be found in Table 3.2.

Proxy Quantities

We define a "proxy quantity" as a quantity that is not present in the prompt of the task, but brought in by the interviewee to replace another. We observed that it was often done to make the task easier to think about. The use of proxy quantities is characterized by an

Reasoning Device	Related MA	Description	Example
Proxy Quantity	MA 1	A quantity used to replace another to make the task easier to think about.	"It's going at a constant rate. So time maps one-to-one to, uh, distance traveled."
Regions of Consistent Behavior	MA 2	Separating a domain into sections that are modeled by the same function.	"So there's a moment when they're doing that loop around Tacoma—looks like half a circle. So that's going to be a constant distance."
Physically Significant Points	MA 3	Plotting only those points that hold physical significance.	"So it's starting and finishing the same distance away...[plots end points]"
Neighborhood Analysis	MA 4	Examining the rate of change of one quantity with respect to another over a "small chunk" around a physically significant point.	"So it's, it's angular velocity has to be faster when it's in the narrow part of the peak...[and] I have more, more, change in the angle for smaller change in the height."
Compiled Models	[none]	The strong association between a particular functional form and a physical context.	"And as you get closer the potential is one over R. So I'm literally just plotting one over R."

Table 3.2: The reasoning devices identified from the physics experts' descriptions of their reasoning as they thought out loud while solving novel graphing tasks. Each reasoning device is associated with—but not the same as—a covariational reasoning mental action identified by mathematics education research (described in Table 2.1).

interplay between MA 1 and codes related to defining, relating, and symbolizing quantities [1]. All physics experts were observed to use proxy quantities.

During our interviews, we found use of proxy quantity occurred most often when participants substituted time for total distance during constant motion tasks. For example, a graduate student working on Going Around Tacoma said:

> "It looks like they're going at a constant speed, so total distance is just going to map to time."

They then proceeded to solve the three tasks in Study 1 by thinking out loud about time instead of total distance while simultaneously labelling the horizontal axis as total distance.

We note that proxy quantities were also observed between quantities that were not related to motion. For example, one graduate student found it easier to think about the number of particles than the intensity during the Intensity task:

> "I want to say the number of particles is proportional to the intensity... I don't know if there's a square in there or not."

They proceeded to solve the task relating number of particles to the depth of the liquid, and then translated the number of particles to intensity, and depth of the liquid to the total distance traveled of the probe.

Regions of Consistent Behavior

The use of "regions of consistent behavior" is defined as dividing a task based on what parts of the domain make sense to be modeled by the same function. We observed that the boundaries between regions were often identified by an abrupt change in motion, and experts tended to focus on modeling one section at a time. This reasoning device is characterized by an interplay between MA 2 and problem approaches such as breaking down the problem into

[1] The related code categories are "Considering Quantity" in Study 1, and "MA 1" and "Quantity" in Studies 2 and 3, as defined in the Appendix for the interested reader

functionally distinct parts [2]. All physics experts were observed to use regions of consistent behavior.

When dividing the tasks into sections in Study 1, all graduate students divided Going Around Tacoma into three sections: traveling from Seattle to the semi-circular path, going around the semi-circular path, and traveling to Portland. For example, one graduate student says upon watching the animation:

> "So just from the picture, I'm just going to split it up into each [of the] segments just because it's easy to find the distance from two lines, and then just from half of a circle."

This observation was consistent across Studies 2 and 3 for the analogous tasks, Gravitation and Electric Charge. Every graduate student in Study 1 also divided Square Track into two sections: going towards the wall and going away from the wall. They all recognized that the turning behavior at the midpoints to the wall did not affect the distance of the cart to the wall.

Experts did not necessarily consider the sections "in order." Rather, they were likely to model each section independently, in whichever order made the most sense to them. For example, many experts began Going Around Tacoma, Gravitation or Potential by first drawing the middle section.

Physically Significant Points

The use of "physically significant points" is defined as choosing and plotting a small number points that hold physical meaning. For these tasks, this physical meaning might represent bounds of a quantity (the maximum or minimum) or important inflection points (e.g., the sides of the Ferris wheel). We observed that experts often chose physically significant points as a first step towards modeling the tasks. This reasoning device is characterized by an

[2]The related code categories are "Approach to the Problem" and "Trends of Change" in Study 1, and "MA 2" and "Graphing" in Studies 2 and 3, as defined in the Appendix for the interested reader

interplay between MA 3 and behaviors such as labeling points on a graph or diagram [3]. All physics experts were observed to use physically significant points.

For example, one graduate student began Electric Charge by choosing some points:

> "So I start at some small positive, go to some bigger positive, stay there, and then go back to my initial point."

They plot four points to represent this story, and then complete their graph saying, "And it'll go like one over R in between." Many of the interviewees were not sure how to begin with the tasks in Studies 2 and 3 that were less familiar, and nearly all of them chose to start by plotting points.

Neighborhood Analysis

"Neighborhood analysis" is defined by creating a smooth graph by choosing physically significant points and drawing small line segments centered around those points that represent the slope of the graph. This was typically followed by connecting these small line segments with a smooth curve and a verbalization of the second derivative. Neighborhood analysis was used when participants did not have a ready-to-apply model from prior physics experience. It is characterized by an interplay between MA 4 and graphing behaviors such as sketching slopes [4]. A majority of physics experts used neighborhood analysis.

For example, a graduate student checking the curve they drew for Drone stated:

> "Because like, the angle is changing really quickly near the middle. You kind of think the derivative should be bigger there. And it sort of is, like, if I look at two, you know, two points around the middle. Like it's changing a lot faster than say two points here [gestures to the far left of their graph]."

---

[3]The related code categories are "Approach to the Problem" in Study 1, and "MA 3" and "Graphing" in Studies 2 and 3, as defined in the Appendix for the interested reader

[4]The related code categories are "Trends of Change" in Study 1, and "MA 4" and "Graphing" in Studies 2 and 3, as defined in the Appendix for the interested reader

This graduate student observed that the angle appears to change faster at the apex of the curve than on the sides, and used that as justification for the slope tangent to the curve at corresponding points on the graph.

Compiled Models

Finally, we define the use of "compiled models" as accessing an already-known model for a particular physical context that carries with it a common functional relationship between two or more quantities. The "$1/R^2$" model for conservative forces is a familiar example to all physicists. Accessing a compiled model provides the reasoner with a wide variety of resources about how two quantities relate to one another. We observed experts use compiled models frequently for tasks that had familiar physical contexts. It is a code category that emerged from data analysis during Study 1, and was included as a code in Studies 2 and 3 (for more information, please see the Appendix). Compiled models is not associated with a mental action besides MA 1 (which we consider trivial in this context), because it involves the application of a model and often does not engage direct covariational reasoning. All experts were observed to use a variety of compiled models.

For example, upon reading the Ferris Wheel task and watching the animation, one graduate student stated:

> "I feel like this is where my like, understanding of trig functions really comes in handy. Because I know this is a circle. And so the height goes like a trig function."

This graduate student then quickly generated a graph using the initial position of the cart to determine the phase of the sine function.

We observed a number of compiled models in various contexts: in constant motion contexts, a linear function that relates distance and time ($d \propto t$); in scattering light contexts, an exponential function that relates intensity and depth ($I \propto e^{-y/y_0}$); and in potential energy contexts, an inverse function that relates potential energy and distance from the source ($U \propto 1/r$).

These associations manifested verbally ("So potential goes like $1/R$"), symbolically, and graphically, suggesting that they hold deep meaning to the user. They were often used in conjunction with proxy quantities to form a final model for the task, or were used to justify the use of a proxy quantity. Because compiled models rely on familiar mathematical models, they are one way that expert physicists were able to arrive at an answer to the tasks without engaging with direct covariational reasoning besides MA 1.

These five reasoning devices often simplified the covariational reasoning tasks, reduced cognitive load, and reduced direct use of covariational reasoning (i.e. explicitly comparing the change of one quantity to another). However, we suggest that these devices appear to be in part habitual, and the avoidance of direct covariational reasoning was likely unintentional.

In addition to simplifying the tasks as part of using reasoning devices, we also observed physics experts simplify tasks explicitly using arguments about symmetry and limiting cases. This result adds to other research in physics education on simplification techniques and limiting cases [158]. Recent work has reported that students who were prompted to evaluate the validity of proposed mathematical expressions for problems in introductory physics used a variety of strategies to do so including the canonically taught strategies of limiting/special cases, unit checking, and considering reasonableness of numbers [6] In a recent study of expert techniques when problem solving across scientific fields, 100% of those interviewed were observed to consider what assumptions and simplifications they could make [111]. Therefore, we focus our analysis on thought processes associated more closely with direct covariational reasoning.

3.4.3 Modes for Graphical Modeling

Our third main finding is that physics experts were likely to engage in one of three modes when graphically modeling the relationship between two quantities. These modes are larger patterns of reasoning that describe the integrated ways in which experts were engaged with covariational reasoning as described by the mathematics education research frameworks (Ta-

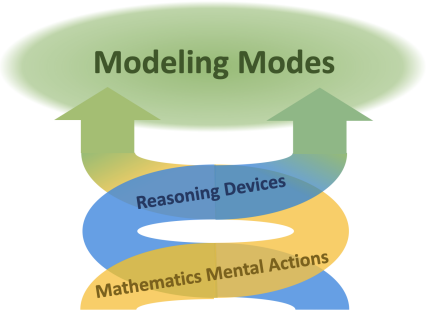


Figure 3.5: The modes identified in our data are representative of a combination of covariational reasoning mental actions as described by mathematics education research and the reasoning devices we identified in Section 3.4.2.

ble 2.1) and the reasoning devices described in Sec. 3.4.2 (Table 3.2). This relationship is shown in Figure 3.5. In this section, we describe how these modes were identified, how they are defined, and how they relate to one another.

During our examination of the participant timelines in Studies 2 and 3, we sought patterns between the code categories. A striking pattern emerged: there was a consistent interplay between the code categories "Function" and "Mental Action," as shown in Fig. 3.6. The code category "Function" describes how experts reasoned using functional forms (i.e. linear, quadratic, exponential, etc.); it includes applying a Compiled Model and using Neighborhood Analysis (as described in Table 3.2). The code category "Mental Actions" consists of MA 1-4 from the mathematics covariational reasoning framework [5]. The interviewees tended to go between these two code categories while modeling—the categories rarely appeared at the same time, and physics experts were likely to use one and then the other back-to-back. The data clustered to three modes of interplay between Function and Mental Actions, which we call Function Knowing, Choosing, and Generating.

Here we use *function* to refer to the functional form of a relationship between two quantities that describes how two quantities change together. This relationship can be expressed both graphically and symbolically. For example, a quadratic function can be represented

[5]The code categories are described in more detail in the Appendix for the interested reader

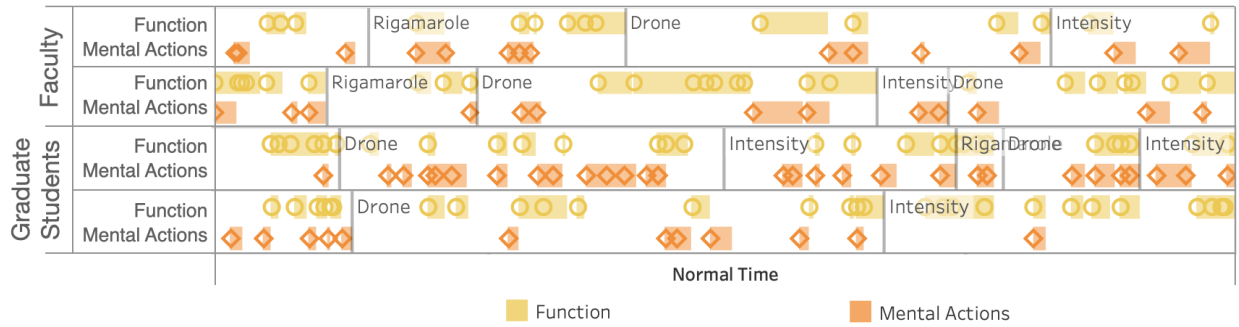


Figure 3.6: Four examples of the interplay between Function and Mental Actions across Studies 2 and 3. These examples were chosen because they are illustrative of the trends we saw for both faculty and graduate students. The horizontal axis represents time on task, and is normalized across participants.

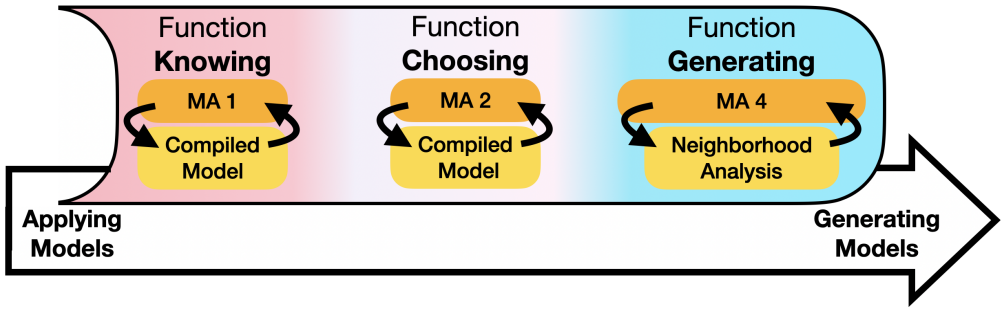


Figure 3.7: This figure depicts the modes we observed experts engage in when modeling, and the interplay between the mathematics mental actions and reasoning devices for each mode. The modes are shown to lie along a spectrum from Applying to Generating Models.

symbolically $y \propto x^2$ and as a parabolic shape in which the dependent quantity varies smoothly in proportion to the independent quantity squared. For the purposes of this work, we consider both of these to be "reasoning about function." Across tasks and expert populations, we observed physicists use the phrase "goes as" or "goes like" to refer to such a relationship (ex: "the height goes like a sine function"). We view the use of the phrase "goes as" or "goes like" to be a clue that someone is engaged in reasoning about a functional relationship between two quantities [171].

Defining Function Knowing, Choosing, and Generating

In this section, we define the three modes, Function Knowing, Choosing, and Generating, and present evidence from across the three studies. We note that the distinctions between Function Knowing, Choosing and Generating are not hard and fast; there were many instances where the research team held discussions about whether a particular portion of transcript was better characterized as one or another. We suggest that these modes describe a spectrum of approaches to modeling from recognizing a well-modeled physical situation towards generating a model where one doesn't exist. This is represented by the arrow in Fig. 3.7, which goes from Applying Models to Generating Models.

Function Knowing is characterized by accessing a compiled model and applying it to the context using important, physically relevant points (such as those described by the initial conditions of the task). Function Knowing often involves the use of proxy quantities to make sense of the accessed model in the context of the task. It is described by an interplay between MA 1 and the use of compiled models (Fig. 3.7).

For example, upon seeing the Ferris Wheel task, many of our interviewees immediately identified the context as circular motion, and that they would need to use a trigonometric function. Most then reasoned about how to "map" between time and total distance to make sense of their initially accessed model that related height and time.

Function Choosing is characterized by using the general trend between physically significant points to choose between a handful of compiled models. It is described by an interplay between MA 2 and Compiled Models (Fig. 3.7).

For example, during Intensity, many experts reasoned that the lowest intensity was at zero distance traveled and the highest at maximum distance traveled. Then interviewees chose how best to connect the points by using the general trend between them and by reasoning about which functions made sense in the physical context—e.g. linear or exponential.

Function Generating is characterized by using Neighborhood Analysis: determining the rate of change at a small number of physically significant points and then plotting the

Reasoning Mode	Item	Example
Function Knowing Quickly associating a physical context with a known mathematical model.	Gravitation	"I want to talk about the gravitational potential energy of the entire system. Okay, so gravitational potential energy goes as $1/R$, one over the total radius to the system."
	Ferris Wheel	"So, the total distance traveled is proportional to the time because it's moving at constant speed."
Function Choosing Plotting physically significant points and using the trend between them based on the physical context to choose a function.	Intensity	"And so we expect that at the maximum distance traveled, the intensity will be greatest. And as light is absorbed, it will get less great...I have no idea. I would expect, and it depends if its in the linear regime, how exactly it will go. But naively, I'd expect that it'd be roughly linear."
	Intensity	"So at the top, you'll get the maximum intensity here and it will drop—it will drop according to some some curve that depended on this. It could be an exponent—and likely an exponential. It really would—should be, uh—should be an exponential fall because the differential probability of having less light drives the thing."
Function Generating Connecting physically significant points with a smooth curve using the rate of change as a guide.	Drone	"The top—basically, the height changes the most slowly, and then the bottom...and, well, as you increase the angles, the height goes—changes faster. So you should have a similar trajectory as you would see in the picture. But how exactly it is, is a different question. I think you can't solve this question quantitatively, but qualitatively you can draw a curve like this."

Table 3.3: The ways in which experts engaged with mathematical modeling while using covariational reasoning during graphing tasks.

points along with small line segments that represent the slope of the curve at those points. The points are then connected with a smooth curve using the line segments as a guide. Some participants would name the function that looked like the curve they drew, and others simply ended the task with the graph. It is described by an interplay between MA 4 and Neighborhood Analysis (Fig. 3.7).

Examples of each pattern are shown in Table 3.3. These examples also help to demonstrate the distinctions between the three approaches. In the examples of Function Knowing, both speakers have a model ready at hand, and spend their time deciding how to apply it. In the examples of Function Choosing, the speakers use the general trend between points to choose between a handful of possible models. Finally, in the example for Function Generating, the speaker does not identify a possible function and instead is guided by their understanding of the first and second derivative. We note these approaches are emergent from the data and are therefore necessarily reliant on the ways the experts chose to talk about how they were reasoning. However, they do give us some insights into common approaches expert physicists took when solving graphical covariational reasoning tasks.

Across all three studies, Function Generating was consistently the last-resort option: experts were likely to begin with Function Knowing or Choosing and finally engage in Function Generating only if Function Knowing or Choosing had failed. Figure 3.7 illustrates the spectrum we observed from applying models with Function Knowing to developing new models with Function Generating. The arrow from Applying to Generating Models is inspired by Tall's 2008 diagram of proceptual reasoning where it represents increasing sophistication [135]. In Figure 3.7, we suggest the arrow is illustrative of increasing cognitive load and the observed reasoning progression.

Evidence of Function Knowing, Choosing and Generating from the Studies

We predominately observed Function Knowing and Choosing during Study 1. For Going Around Tacoma and Square Track, every expert associated the constant motion context with a linear relationship between the distance asked for in the task and the total distance

traveled. Therefore, the time spent on these tasks was focused on connecting physically significant points with straight lines.

During Ferris Wheel, the majority of the graduate students held strong associations between the circular motion context and trigonometric functions. They quickly identified that they needed a sine function, and used initial conditions to identify the phase and the amplitude:

> "So our height is just going to be kind of going sinusoidal...[and the cart] started out at highest point [draws curve]."

Function Generating was used during Ferris Wheel by three of the ten graduate students in Study 1. These participants did not demonstrate an activated association between circular motion and trigonometric functions, and instead examined the rate of change at particular points in order to determine the graph. For example, one graduate student stated:

> "So those — kind of — be some kind of curvature to- to this. So towards the halfway point, when it's almost, when it's halfway to the bottom, the speed should be the greatest. And then otherwise, it should be like — should kind of have a lower slope....
>
> So low to high slope back to low slope."

This student drew small line segments to guide their reasoning, and then connected them with a smooth line to form their final graph (Fig. 3.8).

Studies 2 and 3 were designed to probe expert reasoning outside of constant motion (and circular motion) contexts. Two of the tasks, Gravitation (and Electric Charge) and Intensity required physics knowledge to solve, and therefore necessarily required either Function Knowing or Choosing. However, these contexts were less familiar to most participants and we observed that graduate students were likely to toss out several different models while engaged in Function Choosing. For example, one graduate student stated:

> "Not going to use it, but intensity is proportional to the magnitude of the field squared.

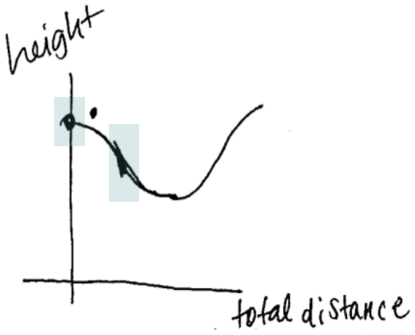


Figure 3.8: A graduate student's work from Ferris Wheel that demonstrates Function Generating reasoning; the point and line segment they used are highlighted in green.

> Yeah, I'm not going to use that. Because...no, I mean, never mind, I can't use that. That is what that's what intensity is. But I'm not going to calculate....
>
> Should I consider the situation...exponential from? I guess that's just me reaching for a model. That happens a lot. Exponential wouldn't make sense. That's like skin depth penetration. That would be like a beam hitting a conductor."

This line of reasoning was ultimately productive for the graduate student, as they were able to get some initial compiled models out of the way by realizing that the models did not apply to the question being asked. However, they also distracted themselves from the correct answer (exponential) with other content knowledge.

This example is provided to show the range of Function Choosing—for some, it was a process of choosing between compiled models that had physical meaning in the context of the task. For others, it looked like coming up with several possible models from physics that are related to the quantities involved, and working through which were relevant to the prompt. Faculty were more likely to engage in the former approach, and early career graduate students were more likely to use the latter.

Finally, Drone prompted participants to engage in Function Generating. This may reflect

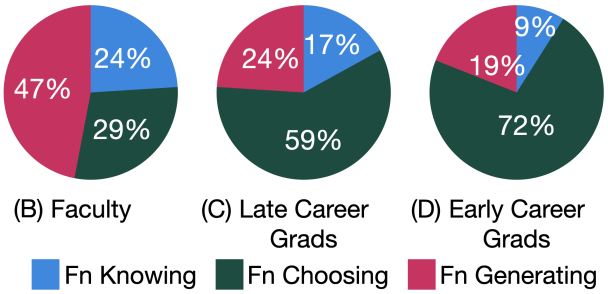


Figure 3.9: The time participants in Study 2 and 3 spent Function Knowing, Function Choosing, or Function Generating as a percentage of time spent engaged in function reasoning.

the fact that deriving an analytical solution is challenging and, as one faculty member said, "messy." Nearly every expert solved Drone by engaging with Function Generating, either as a check on an initially drawn curve or to develop a curve.

Across Studies 2 and 3, we found the graduate students were most likely to engage in Function Choosing, and faculty were most likely to engage in Function Knowing or Function Generating (see Fig. 3.9).

3.5 Discussion

This work seeks to characterize some ways that physics experts reason covariationally while solving novel graphing tasks. Research Question 1 asks "In what ways are the covariational reasoning frameworks developed by mathematics education researchers productive for describing expert physicists' modeling in novel graphing tasks?" We found that some of the ways mathematics covariational reasoning has been described were consistent with physics experts' reasoning, and some were not. Research Question 2 asks "What patterns do we observe in the reasoning of physics experts as they solve novel graphing tasks?" We found five reasoning devices and three modeling modes that were used consistently by all experts while reasoning covariationally. In this section, we describe how these findings relate to current literature and present a hypothesis for how *physics* covariational reasoning is distinct from that in mathematics.

3.5.1 Comparing Mathematics and Physics Expert Reasoning

Our original study—Study 1—was designed as a replication study of Hobson and Moore's 2017 investigation into the covariational reasoning of expert mathematicians. In this section we examine Research Question 1 by considering how our findings are similar to and different from those of Hobson and Moore.

Hobson and Moore reported that mathematics graduate students:

(1) Directly covaried the quantities asked for in the task (in some cases, with prompting from the interviewer), and

(2) Equally divided the domain of the task, or large sections of the motion, into small sections to compare the change in the two quantities and then mapped these equal segments to their graphs [65].

In comparison to (1), we observed that the physics graduate students from Study 1 all used time as a proxy quantity for total distance. Hobson and Moore report that, after prompting from the interviewer, the graduate students they interviewed articulated a direct covariational relationship between the two distances in their equivalent of Going Around Tacoma. For example, they report that one student stated, "for every one mile you move towards the total distance towards [Portland] you're still getting that one mile closer to [Tacoma]" [65]. In contrast, the physics graduate students working on Going Around Tacoma simply stated that because the objects moved at constant speed, distance "maps" to time. One of the graduate students went so far as to create a graph with respect to time. Upon realizing their mistake, they simply crossed out "time" and wrote "total distance" on the horizontal axis. No physics graduate students drew a direct connection between the two distance quantities while thinking aloud or during questioning by the interviewer. We suggest that this reflects how ingrained constant motion contexts are for physics experts, and perhaps also an indication of how ubiquitous proxy quantities are in reasoning about how quantities are related in physics.

We also found that physics graduate students were highly likely to section the domain of a task only if there were distinct regions of consistent behavior. Within a region of consistent behavior, the graduate students either applied an already-known model or used the rate of change to generate a graph. Rather than relate equal portions of change to gather points on the graph and justify a non-linear increase in one quantity for an equal increase in another, the physics graduate students considered nearly instantaneous rates of change centered on particular points and assumed a smooth connection between them.

We recognize that Hobson and Moore were not attempting to characterize all of expert mathematics covariational reasoning; nor are we attempting to describe all of expert physics covariational reasoning. Rather, we take this replication study as an opportunity to consider how practices in the two fields may differ when reasoning about change. Recognizing the differences in how experts reason covariationally may be informative for instructors to reflect on their own reasoning about function and rates of change as it compares to what kinds of instruction their students may have received.

3.5.2 The Emergent Reasoning Devices and their Relationship to the Mathematics Mental Actions

During our analysis across all three studies, we identified several reasoning devices that accompany covariational reasoning used by physics experts. We see these devices as closely related to many of the direct covariational reasoning mental actions defined by mathematics education researchers [28, 141] (towards Research Question 1). There are also ways that these devices are distinct, and we suggest these devices contribute to patterns of expert *physics* covariational reasoning when solving novel graphing tasks (towards Research Question 2).

Proxy quantities were consistently used by physics experts across most of the tasks. In particular, we saw that proxy quantities allowed expert physicists to use models they already knew efficiently and "map" the familiar model to the contexts and quantities the task asked for. We propose proxy quantities as an extension of MA 1 that is particular to physics covariational reasoning, as it is a tool that we observed physics experts use to replace one

quantity for another in order to make the task easier to think about.

This reasoning device is reminiscent of "simultaneous covariation" in mathematics education research [70], which has been characterized by considering how two quantities change independently with respect to time and then comparing the independent changes of one to coincident changes in the other. While a small number of physics experts described their use of proxy quantities as "parameterization," most simply talked about "replacing" distance with time. Further, we observed physics experts engaged in using proxy quantities that were not time, namely, radius for total distance in Gravitation and Electric Charge.

We also observed experts sectioning tasks into regions of consistent behavior, and using neighborhood analysis to examine the rate of change at particular, physically significant points. We see sectioning as an extension of MA 2 particular to physics covariational reasoning, as the experts were seeking sections where the relationship in the trend was constant (i.e. articulating that one quantity increases while the other stays the same for a section of the domain). Similarly, we suggest that neighborhood analysis could be an extension of MA 4 for physics contexts. Thompson and Carlson's 2017 framework defines MA 4 as comparing discrete change in two quantities. The next step (MA 5) is smooth continuous variation, in which the person views the quantities changing smoothly together. We suggest that Neighborhood Analysis sits between these two mental actions; physics experts had in mind that the quantities changed smoothly, but used small, nearly infinitesimal, chunks of discrete change modeled at particular points to guide their smooth curve.

Finally, we suggest there are many differences in the goals of physics and mathematics that might lead to these different approaches in covariational reasoning. As mathematical modeling is often one step towards a larger goal in physics, a degree of speed and approximation is expected. We observed many experts use simplification methods—making assumptions about the task, only focusing on particular points, using limiting cases to constrain the context, and using symmetry to get the second half of a task immediately—to great effect seemingly in order to minimize the effort they needed to extend. It could be that the epistemological approach of finding a model that is, as Albert Einstein said, ""As simple as

possible, but not simpler," is a dominant feature of modeling in physics.

3.5.3 Modes Associated with Modeling using Physics Covariational Reasoning

An important pattern in how experts were reasoning about function while mathematizing during the graphing tasks emerged from the data. We observed three categories of approaches to reasoning covariationally, both symbolically and graphically, which we call Function Knowing, Function Choosing, and Function Generating. In this section, we propose a hypothesis for how these approaches are used and describe how they represent some important ways that expert physicists reason covariationally.

We observed that experts were likely to begin with Function Knowing (if accessible), then work through Function Choosing, if more than one compiled model surfaced. Function Generating was consistently the last-resort option, and only attempted once Function Knowing or Choosing was eliminated. We summarize how Function Knowing, Choosing, and Generating relate in Figure 3.7. From our findings, we described the arrow as descriptive of the observed reasoning progression. We also suggest that it may be related to increasing experience generating models, as we saw that faculty members were more likely to engage in Function Generating than either group of graduate students (see Fig. 3.9).

In our model, Function Knowing, Choosing, and Generating are characterized by an interplay between the direct covariational mental actions defined by mathematics education researchers [28, 141] (Table 2.1) and the reasoning devices described in Table 3.2. The cyclic representation between mental actions and the devices compiled models and neighborhood analysis represent the ways in which Function Knowing, Choosing and Generating are intimately connected to mental actions and go beyond how the mental actions are described in mathematics education research (see Fig. 3.5 and 3.7). In particular, Function Choosing is an interplay between using MA 1 (relating quantity) and accessing a compiled model. Function Knowing is an interplay between using MA 2 (gross coordination) and accessing one or more compiled models. Function Generating is an interplay between MA 4 (chunky continuous) and neighborhood analysis. Function Choosing, Knowing, and Generating are

all founded on reasoning about quantities in the context of the physical world and their change with respect to one another. Thus, we propose that Function Choosing, Knowing, and Generating are three patterns associated with expert *physics* covariational reasoning.

3.5.4 Comparing Faculty and Early Career Graduates

In Function Knowing, Choosing, and Generating, early career graduate students were more likely to stay in the Function Choosing space. Faculty were more likely to use Function Generating overall, but were also the most likely to use Function Knowing of the three populations.

Bing and Redish describe a "journeyman" level that is situated between expert and novice. Journeymen are characterized as "having sufficient skills that they can no longer be considered novices, but not yet had sufficient experience with sophisticated problem solving (and research) to be considered experts" [14]. We suggest that, for early career graduate students, rigorous, mathematically complex problem solving is the norm. It might be that they are so continually thrust into a mindset of complex textbook-like physics problems that they are less likely to solve the tasks we provided with a generative frame of mind. As early graduate students haven't had much experience in research, our data suggest that engaging in the theoretical aspects of research may foster comfort with modeling quickly and making simplifying assumptions. It may also be that "sophisticated problem solving" looks different from the kind of highly complex tasks early career graduate students are asked to engage in during coursework. We suggest that our data for early career graduate students are better described by less experience in research and being in a "test-taking" mindset of coursework in graduate school, but more research is needed to fully investigate these ideas.

3.6 Conclusion

The mathematics education research community has demonstrated the importance of covariational reasoning for understanding concepts in pre-calculus and calculus. Recognition of the importance of covariational reasoning led to the development of a framework of covariation

mental actions, which effectively operationalizes covariational reasoning. Mathematics education researchers have found that the covariational reasoning frameworks in mathematics are helpful when investigating student reasoning.

Though prior work suggests that physics experts' covariational reasoning "looks different" than that described by mathematics education researchers, it also suggests that covariational reasoning is essential for developing and understanding quantitative models in physics. The studies described in this chapter were motivated by a desire to understand how physics experts use covariational reasoning. We examine the reasoning of physics experts engaged in graphing tasks through the lens of the mathematics covariation mental actions framework. Our investigation indicates that the covariational reasoning mental actions described by mathematics education researchers are part of productive quantitative modeling for physics experts across varied physics contexts.

However, we also find that physics experts use a number of other devices, not described by the mathematics covariational reasoning framework, to coordinate two varying quantities. These devices are used alongside the mental actions described in mathematics education research to form patterns in physics experts' reasoning that we identify as modeling modes. We observed these modes across a range of physics contexts and in a range of physics experts. The modes rely not only on direct covariation, but also on facility with the underlying mathematics and physics content. In addition, we find that for physics experts, direct covariational reasoning is intertwined with sensemaking about the physical world and the relevant physics quantities. These results suggest that the covariational reasoning frameworks developed by mathematics education researchers do not fully operationalize *physics* covariational reasoning.

Covariational reasoning has been defined by mathematics education researchers as "the cognitive activities involved in coordinating two varying quantities while attending to the ways in which they change in relation to each other" [28]. We suggest that this definition is productive for physics covariational reasoning as well, with the understanding that for physics experts there exist a number of "cognitive activities" beyond the mental actions described

by mathematics education researchers. We contend that the set of devices and modeling modes described in this chapter, along with the mental actions described by mathematics education researchers, characterizes physics covariational reasoning.

The work described in this chapter provides a foundation for operationalizing physics covariational reasoning. As a next step, we are using the data collected for these studies to investigate how physics experts use sensemaking about the physical world to guide their use of mathematics during covariational reasoning. We are also working to characterize the understanding of the mathematical foundations and physics quantities that seem to be necessary for expert-like physics covariational reasoning. These findings will be synthesized into a framework of covariational reasoning in physics (CoRP). Our hope is that the CoRP framework will not only operationalize physics covariational reasoning but will also be productive for characterizing physics students covariational reasoning.

Chapter 4

EMPIRICAL EVIDENCE FOR THE BLENDED NATURE OF PHYSICS AND MATHEMATICS

4.1 Introduction

Instruction about modeling in mathematics courses often includes separating "mathematical" reasoning from reasoning about the real world. This approach suggests that students coming into physics courses may be primed to reason about mathematics and the real world separately; evidence in physics education research supports this assertion [10, 136]. A growing body of evidence in physics and mathematics education research suggests that this separation is not representative of how those experienced in physics reason [53, 7, 37, 128]. We suggest that this separation may also not be representative of how the experts who design and teach introductory physics courses develop and make sense of models.

If instructors are modeling in physics in ways that are inconsistent with students' prerequisite mathematics courses, both students and instructors may be frustrated with the learning outcomes. Understanding the interface between expert modeling in introductory physics and students' preparation from prior mathematics courses can help improve teaching and learning of physics for both students and instructors. To better understand how well the frameworks common to mathematics instruction align with the ways that physics instructors approach modeling, we sought evidence of whether physics experts reasoned "purely mathematically"—that is, in a context-free way–while developing graphical models.

In this study, we reexamined the data presented in Chapter 3 through the lens of mathematical reasoning. It was quickly evident that physical reasoning was ubiquitous and inseparable from mathematical reasoning in expert statements. This finding led to a shift of the research focus from characterizing a pure math space to a research question that represents

the core of this chapter: "What are some features of the mathematical reasoning used by physics experts when modeling graphically?"

The results we present contribute to the growing body of knowledge about how physics reasoning interacts with mathematical reasoning. Examining the reasoning of physics experts provides researchers with data from one end of the expert-novice continuum, and may help physics instructors recognize their own reasoning patterns, assumptions of comprehension, and emergent expertise in their own students.

4.2 Background

There is a growing consensus that quantitative reasoning in pure math contexts is different from that in physics contexts [130, 84, 67, 159, 116, 25, 11, 151, 128]. Physics education researchers have proposed multiple characterizations of this difference: Redish described physics as a linguistic dialect distinct from the language of mathematics [116]; Caballero pointed to how the operationalization of mathematical tools is often different in physics than in mathematics [25]; and White Brahmia emphasized the centrality of *quantification* to mathematical modeling in physics contexts, where quantification refers to characterizing a feature of a system with a quantity—represented by a letter (e.g., x, v, a)—that includes a unit of measure and typically a sign. [139, 39, 161, 160].

Connecting a quantity's physical meaning to its mathematical representation is an important part of developing and making sense of mathematical models. Reasoning about quantities in the context of mathematical models has been shown to be challenging for many students in introductory physics contexts and beyond, in part because the link between the meaning and representation is not always made clear [103, 160, 166, 39]. Bajracharya and Thompson reported that associating the area under the curve with a particular quantity was an essential skill for success in solving particular graphical analysis tasks [11]. The authors argued that successful students were able to connect the ideas of area under the curve and integration, demonstrating both procedural and conceptual understanding of the underlying mathematics, while also relating the mathematics to its physical meaning. This

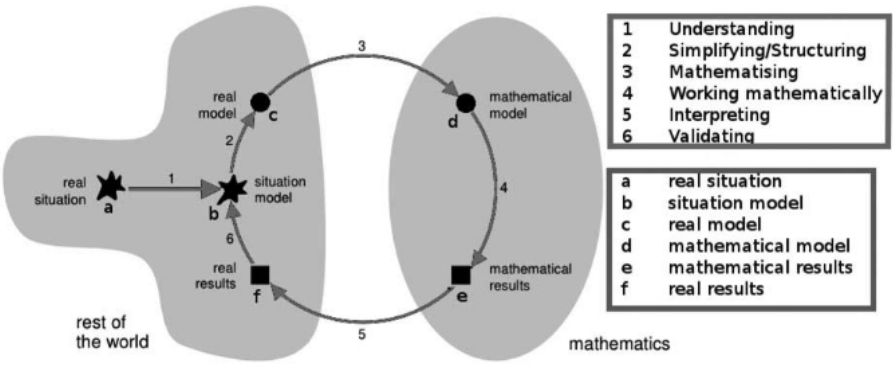


Figure 4.1: Czocher's redraft of Blum and Leiß's modeling cycle [16, 37]. The original Blum and Leißcycle is reproduced in Fig. 2.3.

finding is aligned with research in physics education that facility with multiple mathematical representations is a key element of making sense of the meaning of mathematical models [64, 152, 82, 122, 20].

A common way to represent the sequence of moves in modeling is as a "modeling cycle" in which quantification is an early step, and *validation*—interpreting a mathematical model within the physical context—is the last (see the example from Blum and Leiß in Figure 4.1). Notably, the mathematics space between quantification and validation does not involve consideration of the physical context, but instead represents mathematical manipulation in absence of attention to physical meaning. Similar cycles with separated mathematics and real world spaces have been proposed both in physics and in mathematics education research [16, 115, 117, 148, 167, 89]. While recent education research has moved toward a more blended perspective of mathematics and physics reasoning, these cyclic representations commonly frame modeling education in high school mathematics courses and thereby influence students' perceptions of mathematical modeling before coming to college [149, 57].

Research has found little evidence of clearly separated mathematical and real worlds in quantitative modeling; moreover, the progression of actual modeling may be messier than the published cycles would suggest [53, 7]. In a recent study, Czocher conducted interviews

throughout an academic term of four engineering majors enrolled in a differential equations course [37]. In each interview, the students were observed solving problems in physics contexts that required generating mathematical descriptions from a variety of branches of mathematics, including differential equations. The author described a much finer-grained blending of mathematical reasoning and physical sense-making than is represented in *a priori* "modeling cycles." Specifically, Czocher notes that "there are transitions that appear out-of-order. This was largely because three of the modeling transitions (understanding, simplifying/structuring, and validating) appeared early and often throughout the students' modeling processes" [37]. Serbin and Wawro examined junior and senior physics students' reasoning in quantum mechanics through the lens of Uhden et al.'s modeling cycle, and similarly found that successful modeling patterns did not support the separate math and physics reasoning spaces predicted in the theories of how students generate models. The authors state "when students draw on their mathematical knowledge... it relies on their understanding of how the mathematical concepts and physics concepts are intertwined" [128].

The results of Czocher, and Serbin and Wawro suggest that modeling cycles that maintain separate math and physics spaces may require revision to reflect a tighter blend between physics-like and math-like reasoning [37, 128]. Typically, conceptual blending is used in physics education research to describe a fine-grained approach where particular utterances, equations, or terms are separated into mathematics or physics "input spaces" by the researcher [52, 13, 67, 68, 137, 126, 150, 98]. These input mental spaces are described as subconsciously blended into one idea by the speaker. The blended space may also include utterances that are inseparable by the researchers. Increasingly, physics education researchers are using the term more broadly to refer to the blended nature of mathematical structure and physical meaning [166, 103, 44, 86].

There remains, however, a question in physics about the prevalence of blended reasoning as compared to "purely mathematical" reasoning—i.e. mathematics free of contextual meaning—and what moving between these reasoning modes might look like for physics experts. The research we present here focuses on expert modeling in the contexts of graphing

tasks, adding to current findings associated with the blended mental space between physics and math in modeling.

4.3 Methods

In this section, we describe the methods used to conduct and analyze the think-aloud interviews that were part of our investigation into how expert physicists reason quantitatively during graphing tasks [169].

4.3.1 Data Collection

We report on results from 15 think-aloud, individual interviews that were conducted with 10 graduate students and five faculty members at an R1 university in the Pacific Northwest. Of the 15 experts, six were female or non-binary identifying and nine were male identifying. The pool of experts also spanned a variety of subfields, and included both experimentalists and theorists. An effort was made to include diversity across race and country of origin. Participants were solicited by email, and offered a small gift card to a local coffee shop as a thank you for their time and effort.

The graduate student interview transcripts were divided into two groups of five for data analysis: early career graduate students who were enrolled in required graduate courses, and late career graduate students who were no longer taking courses outside their research area and were fully engaged in research.

One member of the research team conducted the interviews, during which the participants were asked to solve 3–4 graphing tasks. In the analysis described in this chapter, we focus on three tasks that were identical (or isomorphic) across the 15-person expert pool. For the first task, experts were given one of two isomorphic versions: "Electric Charge" or "Gravitation" (Fig. 4.2). For the second and third tasks, all experts were given both "Drone" and "Intensity" (Fig. 4.3). These tasks are the same as those used for data collection in Chapter 3, and repeated here for ease of the reader. All of the tasks prompted the participant to watch an animation and create a graph that related two quantities in the animation.

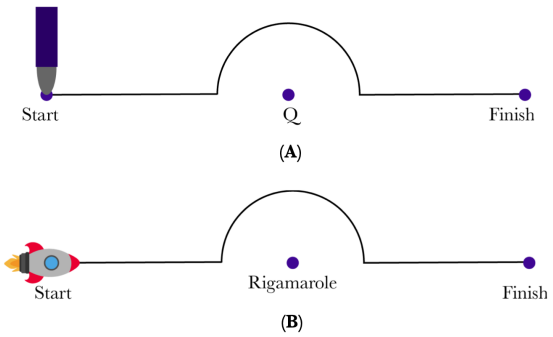


Figure 4.2: Stills from the animations associated with (A) Electric Charge and (B) Gravitation. The tasks prompt participants to create a graph that relates either the electric or gravitational potential and the total distance traveled of the probe or spaceship, as it moves at constant speed from start to finish. This figure is reproduced from Chapter 3 for ease of the reader.

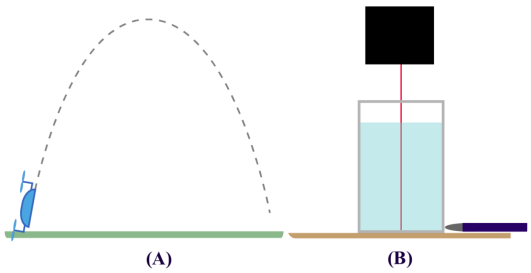


Figure 4.3: Stills from the animations associated with (A) Drone, in which participants were prompted to create a graph that relates the height to the angle of the drone, flying in the arc shown, and (B) Intensity, in which participants were prompted to create a graph that relates the intensity measured by the probe at the side of the liquid column to the total distance traveled of the probe as it moves from the bottom of the column to the top. This figure is reproduced from Chapter 3 for ease of the reader.

Each interview lasted between 25 and 50 minutes, and ended when the participant expressed that they were finished with all the tasks. At the end of the interview, the interviewer asked clarifying questions, if applicable. The interviews were audio recorded; initial transcripts were created automatically using the Otter.ai software program [2], and were subsequently hand-corrected.

4.3.2 Data Analysis

The data were coded using a thematic analysis framework [97]. Thematic analysis emphasizes consistent communication across the research team to ensure reliability of the interpretation of the subjects' statements while recognizing that the research team may bring perspectives to the work based on prior experience with current literature. The practice assumes no *a priori* coding scheme and focuses on grounding the codes in the data through iterative review and discussion. Our interpretation of these data is informed by our prior experiences, established theories of mathematical reasoning [16, 115, 117, 148, 167], mathematics education research testing the cyclic nature of these theories [37, 128], and research in physics education that uses conceptual blending to describe mathematical reasoning in physics contexts [13, 67, 68, 137, 86, 151, 103, 44, 166] as discussed in Section 4.2.

One member of the research team coded the data, using participants' written work and notes from the interviewer to confirm the interpretation of the audio transcripts. Between coding cycles, the researcher conferred with the research team to establish wide-spread agreement.

Once the codebook was stable, the same iterative process was followed to seek themes amongst the codes. Several code categories emerged, three of which we share here:

(1) **Quantity**, which refers to when participants generate, relate, symbolize, or discuss the mathematical structure of quantities which are inherently situated in the real world.

(2) **Task Context**, which refers to when participants made explicit reference to the physical context of the task.

	Code	Description	Example
QUANTITY	Symbolizing	Representing quantity with symbols.	"This is potential, V."
	Mathematical Structure	Choosing a particular structure or discussing why a structure makes sense for a particular quantity.	"So this is only going to be positive."
	Composite Quantity	Defining a quantity in terms of others, or making sense of why a quantity is constructed of others.	"I is gonna be the number sent in times e to the minus y over y_0, where y_0 is a characteristic, sort of, thickness of the material."
	Proxy Quantity	Using one quantity in place of another to make sense of or visualize what is happening.	"So as I go from distance traveled, I'm basically starting at a fixed R_0 and I'm going to a smaller R [from] there."
	Code	Description	Example
TASK CONTEXT	Connection to the Task	Making sense of a quantity or feature of their graph, or drawing a diagram to visualize a quantity in the context of the task.	"And so when they're far apart, there's low gravitational potential energy."
	Code	Description	Example
MODELS	Compiled Model	When participants bring in a previously known model from physics.	"Okay, so I'm going to sort of invoke Bier's law."
	Compiled Trend	When participants refer to a particular function and/or discuss its behavior.	"I'm going to say this is a graph of something like a minus arctan of x."

Table 4.1: Codes associated with the code categories Quantity, Task Context, and Models.

(3) **Models**, which refers to the ways in which participants use previously known physical models or apply their knowledge of particular functions common in physics.

Table 4.1 shows the codes in these categories, along with a short description and example of each code.

The research team took care to allow the codes to emerge from the data. Upon first glance, a code for "physical reasoning" seemed appropriate, however it quickly became clear that the "physical reasoning" code was ubiquitous across the data, rendering it meaningless. The lead coder also felt they could not distinguish between "physical" and "mathematical" reasoning throughout the data, even after lengthy discussion with the research team. Therefore, we instead developed the code "Connection to the Task," and established the boundary that this code only included instances in which experts explicitly relate their reasoning to the task at hand. This distinction removed some of the subjective interpretation associated with deciding what kind of reasoning is physical, and grounded the coding in the statements of the experts. A consequence of this decision is that reasoning about an abstract physical model is not included in the "Task Context" category. For example, one expert stated:

> "If I think about v_x over v_y, v_x is a constant and v_y is linear in t."

This statement is categorized as "Quantity," since the expert is symbolizing velocity quantities, but not "Task Context" since they are reasoning using the kinematic equations in abstract terms. Instead, we applied a separate code, "Compiled Models," to indicate they were using a previously-known model [169]. In contrast, later in their reasoning this expert stated:

> "As we get towards the top...as my height increases, my angle—does it change faster?"

It is clear the expert is considering the context of the task at hand when they talk about "the top" and the statement is categorized as "Task Context."

We created timeline charts to seek patterns between code categories. The timeline charts visually represented segments of transcript for when a participant's reasoning was assigned a particular code category, and the duration for which they were coded in that category. Figure 4.4 shows six representative timelines from our data. We sought patterns in which categories frequently appeared together and frequently appeared apart. After identifying segments of transcript in which a given pattern appeared (e.g., context coded with quantity), we returned to the transcripts for further analysis. We categorized the transcript segments into two groups: (1) those in which the text supported the pattern and (2) those in which the pattern was refuted. For example, for a pattern that appears together in the timelines, we collected all transcript pieces where the code categories coincided, and separately collected all transcript pieces where the code categories had not. We examined these sections of text carefully to find patterns in the ways that experts reasoned in order to characterize, or dismiss, the relationship identified in the timeline.

Finally, we used quantitative measures to compare the frequency of categories appearing separately or together. For example, we calculated the time on task for all utterances where Quantity had been coded (on its own, or with any other category), and the time on task where Quantity and Task Context had been coded together.

4.4 Findings

Our research question for this study focuses on the relationship between physics and mathematics reasoning: "What are some features of the mathematical reasoning used by physics experts when modeling graphically?" Our data revealed that there was almost no time during which physics experts reasoned without reference to the physical world. This is in contrast to what is suggested by the presence of "purely mathematical reasoning" spaces in common modeling cycles. Instead, the distinction between how experts reasoned physically and mathematically lay in whether they reasoned with reference to the physical world abstractly or with reference explicitly to the context of task at hand.

We identify two major features of mathematical reasoning in physics from this work: (1)

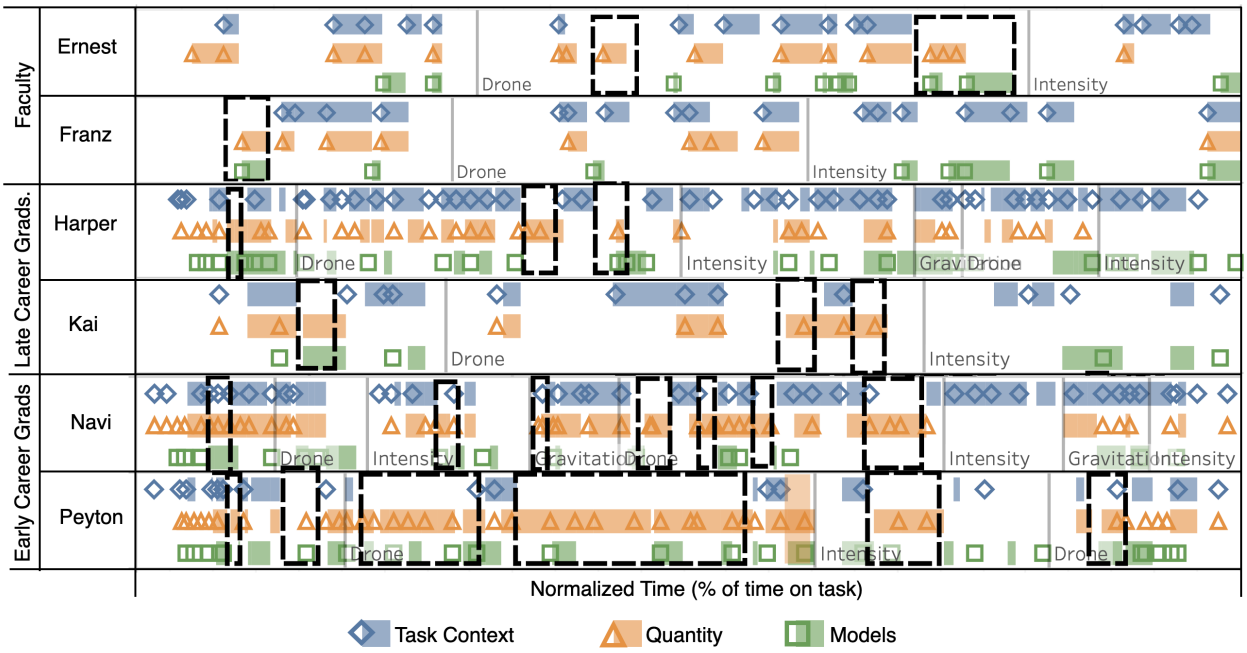


Figure 4.4: A visual representation of when participant statements were assigned the "Quantity," "Task Context," and "Models" code categories. The length of the bars corresponds to the time the participant was engaged in that line of reasoning. The horizontal axis represents time, normalized to the length of each interview. Therefore, it can be thought of as a percentage of their total time on task. The outlined regions are where "Quantity" appears without "Task Context." Gray vertical lines represent when participants switched tasks.

when physics experts reasoned mathematically, they generally did so with explicitly reference to the context of the task, and (2) during the relatively infrequent occasions when experts did not attend explicitly to the task context, physics experts were likely to use other familiar physics models as a starting point to find an approximate, analytical solution. During this time, we observed experts algebraically manipulate those models to see if they could be adjusted to fit the novel context of the task. We name this behavior "mathematical riffing," and characterize it here. We found that early career graduate students spent a larger percent of time-on-task riffing, as they were less likely to reject an unproductive model and thus spent longer trying to make the unproductive model work. In this section, we describe each of these findings in detail.

Blended Mathematics and Physics

A dynamic interplay between behaviors categorized by Quantity codes and Task Context codes emerged across all experts. The timeline charts (Figure 4.4) indicate that behaviors coded as Quantity and Task Context appeared together frequently. The Quantity and Task Context categories were assigned at the same time when the research team could not separate the ideas within a single utterance. For example, one expert working on Gravitation stated:

> "It's got to be negative one over R, because I want it to fall in."

The speaker associated their symbolic expression with what it implies is happening in the physical system—the negative sign is required for the model to describe an attractive gravitational field that the spaceship can "fall in" to.

These kinds of statements were consistent throughout the data; we found that about 3/4 of the time, Quantity codes were assigned for the same utterances as Task Context codes (Fig. 4.5). We interpret this to mean that, for experts engaged in graphing tasks, the majority of their reasoning about symbolic representations is grounded in the physical context of the task they are working on.

We also observed that when the experts were making sense of quantities, the mechanism for how the quantities might affect one another was an essential tool [124, 123]. For example, Gravitation prompted many physics experts to reason about the physical mechanism that causes changes in the spaceship's motion. Uniformly, the experts often produced an already known model of potential energy as a function of distance from the source (either graphically or symbolically), and then used it to create new graphs of potential energy as a function of total distance traveled. From there, several made sense of the new graphs by thinking about what happens to the spaceship's speed as the object gets closer to the source. One faculty member stated,

> "It's going to speed up as it gets zooming into this thing. So I'm going to swing it down here and swing it up there because there's more kinetic energy down here

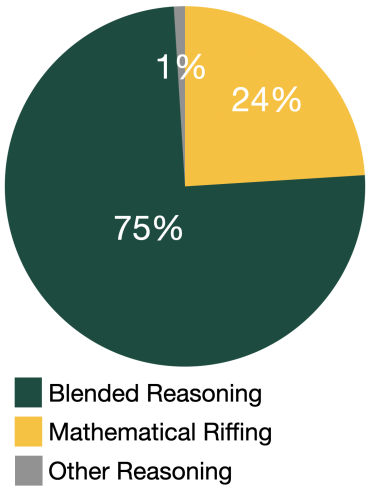


Figure 4.5: Percentage of time spent reasoning in a blended way, which we measure as Quantity and Task Context coded together, abstractly, which we call "mathematical riffing" and is measured as Quantity coded without Task Context, and other reasoning, which is measured as when Quantity is coded without Task Context, but the transcript reveals that the participants are not riffing.

> [referring to the points closest to the planet Rigamarole]."

For some experts, using kinetic energy to reason about their potential graph was somewhat complicated by the animation, which shows the spaceship moving at constant speed. A graduate student reasoned through the mismatch in expectation, stating:

> "There's probably some energy source the shuttle is using when it's flying away, is my guess. And when it's braking. 'Cause I don't know how else it will take that path."

Similarly, in the Intensity task many experts considered the effect of the liquid. Several experts noted that they couldn't imagine a glass of water changing the intensity in a measurable way:

> "I don't know, it's hard to imagine a laser light being that much affected by water, but I guess so. So as soon as you went infinitely far down into a glass,

> there would be so much light scattered before reaching the bottom that there would be no light."

Many used the ocean as an example, preferring to think about a body of water large enough that they could imagine the light diminishing. Of those physics experts who convinced themselves the light may be affected measurably, most then wrestled with whether a linear or exponential function would be more appropriate. This debate was either because they were considering the attenuation of the liquid, or because of self-described inexperience with the context.

These examples illustrate how the physics experts were engaged not just in the technical definition of the quantity—its unit, sign and physical meaning—but also the *physical implication* of its meaning at the same time. The quantities bring to mind both particular functions or contexts and a sense of physicality; the physics experts expressed ideas about what quantities made sense to change and what mechanism would cause the change [1].

"Mathematical Riffing"

When transcript portions are categorized as "Quantity" but not categorized as "Task Context," we typically see physics experts setting aside the context of the task to reason algebraically. However, the algebraic reasoning we observed is not "purely mathematical," that is, we observed it is not devoid of contextual or physical meaning. For example, one expert working on Drone stated:

> "I think all I know is the... x is going to be some $v_0 t$... And v_0 is related to θ_0. So let's write x of t."

This expert makes a series of algebraic moves based on their assumption that the problem can be modeled with a constant horizontal velocity. Their algebraic reasoning carries with it

[1] For more information about how the physics experts reasoned about rates of change and functions in particular, please see our accompanying publication [169].

implicit physical meaning—v_0 represents velocity, and is introduced as an already established fact (despite this being its first occurrence in the interview). This quote is representative of what we observed across the expert pool when we reexamined sections of transcript categorized as "Quantity" and *not* "Task Context": the reasoning seemed to be anchored in already familiar physical models. These moments are outlined in Figure 4.4. The timeline charts illustrate that physics experts are likely to engage with physical models before, after, and often during these periods of mathematical reasoning (as shown by the prevalence of green squares, indicating "Models" near outlined sections of the timelines).

We name this behavior "mathematical riffing": the physics experts are playing around with familiar models to generate a new expression. We use this language as an analogy to the practice of riffing by jazz musicians. To the untrained ear, jazz improvisation may appear to be random. However, playing improvisational jazz is deeply informed by the musician's experience and knowledge. Phrases such as "What happens if..." or "Let's see what we can do with this..." were strong indicators that an expert was engaged in mathematical riffing. During this process, it was common for physics experts to refer back regularly to the task to check that what they were doing made sense—in the timeline charts, this is seen by the Task Context code category appearing in between, or at the ends of, outlined sections. These checks were used to make sense of the model they were generating and to confirm, or reject, an emerging model as productive towards solving the task. To exemplify mathematical riffing, we provide a transcript excerpt from an interview with one of the faculty members whom we call Ernest.

While working on the Drone task, Ernest recognized that the parabolic shape might mean the motion of the drone could be modeled using a quadratic function. He stated:

> "So [the height of the drone] is going to be minus some constant times x squared, which we can work out.
>
> "This is just assuming it's a quadratic, actually, I don't know that for sure.
>
> "Well, we said no air resistance, so we do know that for sure."

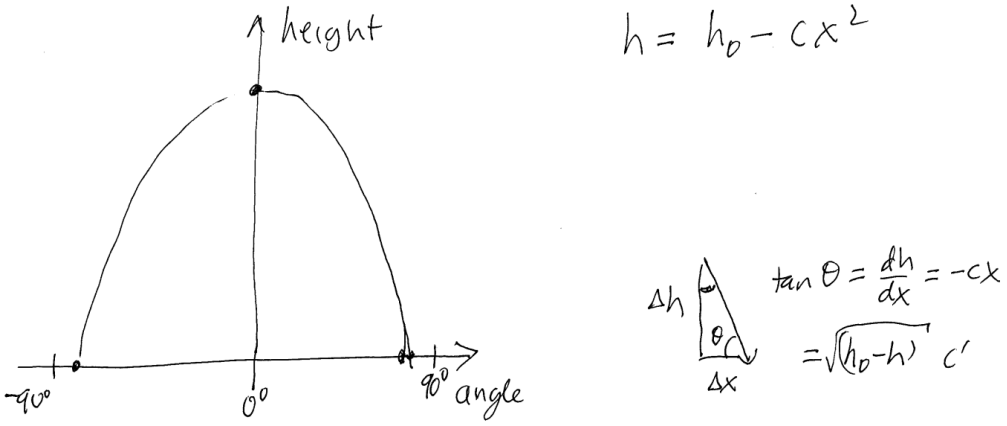


Figure 4.6: Ernest's written work for the Drone item.

Ernest wrote $h = h_o - cx^2$ on his paper to define the height, h, by the horizontal position of the drone, x, and then checked his expression against the context of the task to make sure it made sense (See Fig. 4.6). The reference to the absence of air resistance suggests this model may be linked to a model of constant acceleration for Ernest, but he did not state so explicitly. Later, he worked on defining the angle, θ:

> "The tangent of this angle is going to be given by dh by dx.
>
> "Okay, which is minus c times x.
>
> "And that's in some complicated relation to h, which is given by h_0 minus h, square root, thereof and some factor, which I could work out. So, okay, so this is some messy formula, I'm not going to come up with this very well.
>
> "So we'll just draw a smooth curve, which goes between these two, and we'll call it a day at that point, okay."

Ernest determined a symbolic representation for both h and θ, and attempted to find an expression of h in terms of θ. However, he quickly changed tack when he realized the expression would be complicated and possibly unhelpful towards completing the task. During

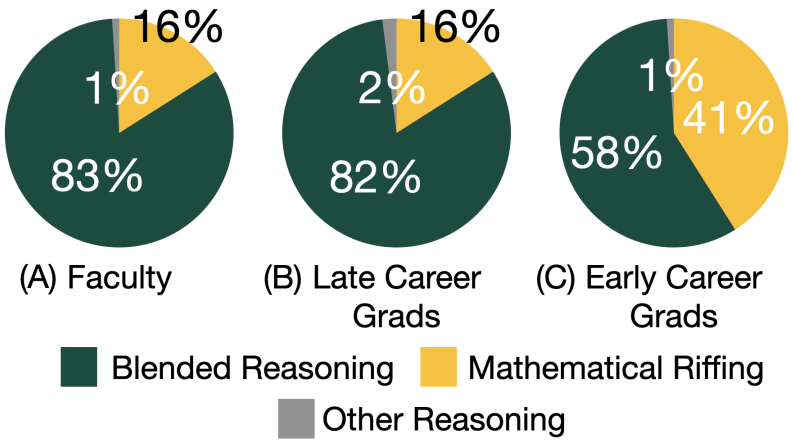


Figure 4.7: Percentage of time spent reasoning: in a blended way, as measured by Task Context and Quantity coded together; mathematically riffing, as measured by when Quantity is coded without Task Context; and Other, as measured by when Quantity is coded without Task Context but the transcript reveals that the reasoning is not fully aligned with our definition of mathematical riffing. Percentages were calculated as percent time on task for each participant, and averaged by population to produce the charts shown.

this time, it was clear that the symbols held physical meaning for Ernest, even when he was disconnected from the task itself.

The majority of the physics experts engaged in mathematical riffing at some time during the interview, suggesting it is a useful tool and a mark of expertise. We observed that productive riffing is characterized by adjusting established physical models to describe a novel context. Equally important to productive riffing is the flexibility to change tack quickly when a model does not appear to work well for a given task.

We also observed unproductive riffing, which is characterized by continuing to persevere with a model that appears unlikely to be useful. As an example, we provide a short excerpt from an early career graduate student, whom we call Peyton. While working on Drone, Peyton spent a long period of time attempting to derive an analytical expression:

> "...So right now I have y is equal to $x \tan\theta$, which is equal to x_0 plus $v_x t \tan\theta$. I don't want to eliminate θ, but I'd like to eliminate time...

> "If I put y over $\tan\theta$ into here... then... have something that's quadratic in both. I don't really like that..."

Peyton continued in this manner for a large percentage of their time on task, as shown by the dashed boxes in Figure 4.4. Throughout their time reasoning algebraically, they seemed unconvinced that they were on the correct path, regularly making statements like "This looks messy..." and "That seems weird." Despite their concern, they continued working with the same model for the majority of their time spent solving Drone.

We observed a difference in prevalence of mathematical riffing across populations. Faculty and late career graduate students were less likely to spend much time riffing, as shown in Figure 4.7, which we attribute to the fact that they were quick to abandon unproductive models. The early career graduate students percentages are dominated by Peyton, who engaged with mathematical riffing far more than their peers and often in unproductive ways. However, even when we removed Peyton from the data set, a trend is apparent: the average percent time on task riffing for the rest of the early career graduate students was 30%. A full timeline chart of each participant can be found in the Appendix for comparison.

4.5 Discussion

In this study we focus on how physics experts engaged with mathematical and physical reasoning during a series of graphing tasks. The tasks in these interviews were designed to probe graphical reasoning, and therefore did not prompt participants to generate symbolic expressions. However, many experts spontaneously chose to develop symbolic representations to help themselves reason. We found that experts engaged in physical sensemaking whether reasoning symbolically or not; physics experts continually maintained a connection to the physical world throughout their interviews. There was little to no evidence that the physics experts were ever reasoning in a purely mathematical space, that is, one in which symbols hold no physical meaning. This finding suggests that mathematical reasoning in the contexts of the chosen tasks almost always contains some reference to the real world for physics

experts.

It was common for the physics experts to reason with specific functions to help them reason through the graphing tasks, which may feel to them like they are doing "pure math." However, we found there was always physical reasoning embedded in their mathematical reasoning—the variables carried physical meaning and experts had expectations about how they would behave as another quantity changed. These periods of algebraic reasoning, what we are calling mathematical riffing, are characterized by physics experts manipulating already familiar physical models, seeing what happens, and either noticing when the model is not fruitful for them and abandoning the line of reasoning, or building on the model. Essential to mathematical riffing is that the expert bases their reasoning on expected common functions in physics and physics models. The expert manipulates these models in a way that may appear inexpert-like to a novice, but is informed by an understanding of what makes sense for the context. This finding suggests that physics experts associate physical meaning with algebraic variables in ways that mathematicians typically don't, and that it is an essential part of mathematical thinking in physics.

Collectively, our data can be interpreted through the lens of a conceptual blend in which mathematical and physical reasoning combine to form a distinct way of thinking. This finding contributes empirical evidence of conceptual blending by successful problem solvers in lower-division contexts to the body of existing literature. It also suggests that models that include separate reasoning spaces for "purely mathematical" and "real" worlds should be reconsidered.

We also found that the physical mechanism by which quantities changed was integral to how physics experts reasoned about the meaning of their representations. Research has shown that using the physical mechanism of a system to make sense of why a quantity may change is an important aspect of conceptual physics reasoning [61, 124, 118, 74]. Euler, Rådahl, and Gregorcic describe how mechanistic reasoning is an important part of students developing explanatory models; research in physics education has demonstrated the value of explanatory models in modeling instruction [49, 51]. This study provides further evidence

that mechanistic reasoning also informs the development of mathematical models by physics experts.

Finally, we observed differences in the effectiveness of mathematical riffing across our expert populations. Figure 4.7 shows that the early career graduate students in our sample spent more time mathematically riffing than more experienced experts, because they were less likely to reject an unproductive line of reasoning quickly. This emergent pattern is aligned with Bing and Redish's description of a journeyman: "that level where students have developed sufficient skills that they can no longer be considered novices but where they have not yet had sufficient experience with sophisticated problem solving (and research) to be considered experts" [14]. Bing and Redish compared novice and journeymen reasoning, and showed that journeymen more quickly recognized their approach was unproductive and were more likely to change tack than novices. Our data represent another section of the spectrum, from journeyman to expert. Together, these results demonstrate that rejecting unproductive models is a hallmark of expertise, and that this skill may be slow to develop across many years of physics experience.

Given the small sample size and the potential for embarrassment while problem solving in an interview, we emphasize that further research is needed here before we strongly claim that there is an "early-career" effect. We cannot ignore the potential influence of imposter syndrome—the participants were aware the study was focused on expert reasoning, and this may have had an impact on how they felt during the interview. Similarly, early career graduate students may be so strongly situated in complex problem solving for which they expect a solution that they may have been more willing to perservere to an exact analytical solution.

In summary, we interpret our findings that, in the context of challenging, novel, classical physics graphing tasks, physics experts continually maintain a tie to the real world, even when reasoning mathematically. This finding is aligned with those of Czocher and Serbin and Wawro [37, 128]. Despite the differences in methodology and context—both Czocher and Serbin and Wawro examined tasks that prompted symbolic answers, interviewed under-

graduate students in engineering and physics respectively, and used *a priori* coding schemes based on established modeling cycles—the results are strikingly similar. These complementary findings do not support the prevalence of separate math and real worlds in the contexts encountered by the vast majority of students who take physics courses, rather, the findings support descriptions of modeling that incorporate a continuous interaction between the mathematical and physical worlds as blended reasoning. Future research might consider revising theoretical models to support expert behaviors of carrying physical models with them into "pure" math spaces.

We encourage instructors to reflect on their thinking in the context of modeling by noticing how and when physical ideas are incorporated, often implicitly, when doing math in physics contexts. Even considerations such as why particular arithmetic operations make sense in a model often have physical foundations (i.e., summation in conservation laws, subtraction and division in rates of change). Acknowledging this subtle expert knowledge might help students recognize what is familiar, and what is new about the modeling that they are learning to do in physics.

4.6 Conclusion

In this chapter, we present empirical evidence of the emergent reasoning of physics experts while they develop graphical models. While there is a spectrum of reasoning from more math-like to more physics-like, there is scant evidence in our data of purely mathematical reasoning. The majority of the time, physics experts' reasoning can be described as a conceptual blend of mathematics and physics context. When experts reasoned mathematically without considering the context of the task at hand, they engaged in mathematical riffing—using abstract physical models as a starting point, adjusting the mathematics to work towards a novel expression, and quickly rejecting models that are not productive. Considering physical mechanisms played an important role in their decision-making.

This study was limited by the context of the tasks, which were all prompting for graphical representations (although several experts spent a good fraction of their time generating and

engaging in symbolic representations) of some aspect of constant speed motion in classical physics contexts. The narrow scope of the contexts correspondingly limits the generalizability of the results; we do not claim that our findings characterize physics reasoning in every context.

Additional research can help span the space of expert mathematical reasoning in physics. For example, we can imagine enriching these findings by replicating this study in more abstract contexts, such as quantum or statistical mechanics. While it was difficult to generate prompts that are actually novel for experts, our findings suggest that the process of examining experts generating models is potentially fruitful for future research.

Basing instruction on an assumption that mathematical spaces and real world spaces exist separately is not consistent with the way that physics experts appear to model. Instruction that implies they are separate spaces might cause confusion for novices as they are learning from experts who largely don't reason this way. We encourage instructors to reflect on their own experiences of "doing math" in physics, and to be aware of implicit assumptions about deeper physical meanings that are nearly always embedded in the mathematical reasoning of introductory physics. In addition, instructors, specifically when rejecting a model, may find it productive to consider ways to make their decision-making process explicit to the students in their physics courses. In general, instructional interventions that help bridge gaps between what experts may consider foundational mathematical reasoning in physics and what novices are prepared for in their prerequisite courses will enrich the learning of physics. Our findings can help the ongoing and future efforts to close these gaps.

Chapter 5

THE COVARIATIONAL REASONING IN PHYSICS FRAMEWORK

5.1 Introduction

In a typical physics lecture, statements such as "...it goes like $1/r$..." or "...they're inversely proportional..." are common, and often considered part of "thinking like a physicist". More specifically, these statements are examples of *covariational reasoning*—considering how changes in one quantity affect another, related quantity. Covariation is central to reasoning in physics, especially reasoning related to quantitative modeling. Given how productive this reasoning is in physics, we would like our students to leave our courses with a facility with and tendency for reasoning covariationally. Unlike well established content-based learning objectives (such as Newton's Laws or conservation of energy), it can be difficult for instructors to know how to teach students to reason this way. Covariational reasoning that can be simply adapted to physics contexts is not currently an outcome of prerequisite mathematics courses.

While many students enter physics courses with some experience of covariational reasoning from previous courses in mathematics, research has demonstrated in that reasoning learned in mathematics courses does not translate directly to physics contexts [116, 84, 67]. Mathematical reasoning doesn't simply "come along for the ride" in physics contexts. Indeed, physics students' covariational reasoning does not improve substantially as a result of instruction in introductory-level physics, even in settings where the instruction is research-based [166]. Moreover, prior work demonstrates that physics experts' covariational reasoning is not fully characterized by covariational reasoning as described in the mathematics education literature (see Chapter 3). This suggests that *physics covariational reasoning* is distinct

from covariational reasoning as taught in mathematics courses. Therefore, in order to design instruction that does improve physics students' covariational reasoning, physics covariational reasoning must first be characterized in an assessable, fine-grained way.

Using prior work by both mathematics and physics education researchers, we developed the Covariational Reasoning in Physics (CoRP) framework to formally characterize physics covariational reasoning. The CoRP framework distinguishes physics covariational reasoning from covariational reasoning as described in the mathematics education literature. In this chapter, we describe the development and relevance of the CoRP framework. We demonstrate its use for instructors and researchers as a tool to analyze students' covariational reasoning, and describe how it can be used to aid in development of instructional interventions to target students' physics covariational reasoning directly.

5.2 Background

Our understanding of physics covariational reasoning is built on work by mathematics education researchers, and informed by recent work in physics education research. In this section, we describe the frameworks of covariational reasoning developed by mathematics education researchers that form the theoretical backbone of our research. We then discuss recent work focused on physics experts' covariational reasoning that provides the experimental basis of the CoRP framework.

5.2.1 Covariational reasoning in mathematics education research

Covariational reasoning has been defined by mathematics education researchers as "the cognitive activities involved in coordinating two varying quantities while attending to the ways in which they change in relation to each other" [28]; that is, it describes reasoning about how changes in one quantity affect changes in another, related quantity. In mathematics education research, covariational reasoning has been studied widely and has been identified as an essential part of reasoning in pre-calculus and calculus [138, 31, 75, 23, 45, 19]. Covariation has been studied in contexts of reasoning about function [36, 29, 99, **?**, 141, 107], as well as

graphing and the use of coordinate systems [95, 157, 72, 65, 24]. It has also been identified as necessary for reasoning about rates of change [139].

In 2002, mathematics education researchers Carlson, Jacobs, Coe, Larsen, and Hue developed frameworks describing hierarchical levels and associated "mental actions" (MA) of covariational reasoning [28], based on studies of undergraduate math students interpreting and creating representations of functions [29]. The covariational reasoning mental actions were designed to allow researchers and educators to assess the level of students' covariational reasoning. Each of the mental actions is associated with specific behaviors related to covariational reasoning. The mental actions range from a recognition that variables are related (MA 1), to considering the specific relationship between the variables, including the rate of change and the rate of the rate of change (MA 5). In 2017, mathematics education researchers released an updated framework of covariational reasoning [141], incorporating research performed subsequent to the development of the original framework. Table 5.1 shows a summary of the relevant aspects of the 2002 and 2017 covariational reasoning frameworks, adapted from Jones [76].

5.2.2 Covariational reasoning in physics education research

In physics education research, reasoning about how two or more quantities change with respect to one another often falls under the names of proportional reasoning or scaling [5, 18, 17, 15]. Proportional reasoning typically refers to directly proportional relationships (i.e. $F \propto a$), and has at times been extended to refer to non-linear relationships (i.e. $U \propto -1/r$). Scaling is often used in geometric contexts; howeve,r it is also used throughout the literature to refer to relating discrete changes of two quantities (i.e. "if I double this quantity, what happens to that quantity?"). In the language of covariational reasoning, we consider proportional reasoning to be *linear* covariation reasoning and scaling to be an instance of *discrete* covariation.

Work in physics education research has demonstrated that the language of covariational reasoning from mathematics education is helpful in analyzing novice and expert work [137,

Label	Mental Action [?, ?]	Brief Description [76]	Example Behavior
MA 1	Recognize Dependence	Identify variables that are dependent	Labeling axes
MA 1.5	Precoordination	Asynchronous changes in variables	Articulating that first, one quantity changes, and then the other changes
MA 2	Gross Coordination	General increase/decrease relationship	Describing that as one quantity increases, another decreases
MA 3	Coordination of Values	Tracking variable's values	Plotting points
MA 4	Chunky Continuous	Values changing in discrete chunks	Articulating that as one quantity doubles, the other triples
MA 5	Smooth Continuous	Continuous, simultaneous changes	Describing that the quantities vary together, smoothly and continuously

Table 5.1: A summary of the covariational reasoning mental actions (MA) frameworks developed by mathematics education researchers [?, ?]. Summary adapted from Jones [76]. This table is repeated for ease of the reader.

46, 104, 90, 151, 172, 170, 133]. Recent work by Zimmerman, Olsho, Loverude, and White Brahmia has sought to explore the extent to which the mathematics covariational reasoning mental actions framework can be used to analyze *physics* covariational reasoning of physics experts engaged in modeling tasks (see Chapters 3 and 4 [169, 168]). The Zimmerman et al. study involved individual, think-aloud interviews with 20 physics experts (graduate students and faculty) engaged in tasks designed to elicit covariational reasoning. The tasks prompted the participants to create a graph that related two quantities. For example, one task depicted a Ferris wheel cart in motion (see Fig. 5.1) and asked participants to relate the height of the cart and its total distance traveled. The overarching results of study were that:

1. Physics experts demonstrate mathematical reasoning that is consistently woven into physical sensemaking of the quantities involved (See Chapter 4).

2. Physics experts engaged in patterns of covariational reasoning and modeling that were not well described by the mathematics covariational reasoning framework summarized in Table 5.1 (see Chapter 3).

The CoRP framework synthesizes the mathematics covariational reasoning frameworks, the recent study of physics experts' covariational reasoning, and other prior work in both mathematics and physics education, to operationalize the use of covariational reasoning in introductory-level physics.

5.3 Characterizing Covariational Reasoning in Physics

In this section, we describe the devopment and meaning of the CoRP framework. In section 5.3.1, we describe foundational work and theoretical underpinnings that have directly influenced the development of the CoRP framework. In section 5.3.2, we describe the steady-state version of the Covariational Reasoning in Physics (CoRP) framework.

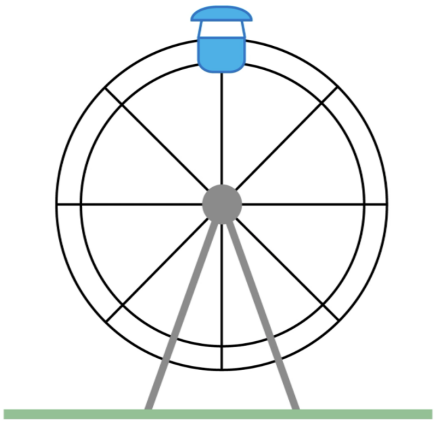


Figure 5.1: A still from a covariational reasoning graphing task that asks experts to create a graph relating the distance traveled by a Ferris wheel cart and the height of the cart.

5.3.1 Underpinnings

Mathematization and Proceptual Understanding

Mathematization in physics is the association of a system or context from the physical world with a mathematical representation (e.g., symbolic or graphical). One aspect of mathematization in physics is *quantification*—the process of conceptualizing a system and a quality of it so that the quality includes a value, a unit of measure and very often a sign.[139, 160]. Research in mathematics and physics education suggests that mathematization and quantification are challenging for students; for example, it has been demonstrated that students are unlikely to think of compound quantities, such as momentum, as quantities unto themselves [138, 153, 147, 163, 160]. In addition, students are likely to come into physics courses with some foundational mathematical resources but may not yet use them productively for mathematization in physics [60, 18, 160, 166, 98]. For example, introductory physics students typically have a mastery with multiplying and dividing numbers; however, they may struggle to conceptualize product quantities (such as momentum and energy) and ratio quantities

(such as velocity and acceleration) as distinct from the quantities that contribute to their calculation [145, 146, 18].

Mathematization, including quantification, depends on connecting mathematical ideas with what they represent in the physical world. Students engaged in modeling tasks have been shown to refer back to the context of the task consistently throughout a productive modeling process as part of making sense of and validating their model [37]. *Proceptual understanding* is defined by Gray and Tall as a combination of *pro*cedural mastery and con*ceptual* understanding [56]. For example, in the context of fractions, "the symbol $\frac{3}{4}$ stands for both the process of division and the concept of fraction"; that is, a student with a proceptual understanding of fractions would move fluidly between the procedure of dividing 3 by 4, and the instantiation of the fraction $\frac{3}{4}$ as a precise quantification of portion. Maintaining a proceptual view of a mathematical representation—consistently making sense of both the mathematical formalism and the physical meaning of the expression—is a key part of modeling, and therefore essential for covariational reasoning.

Mathematics and physics meaning are interwoven in physics. This blended nature of physics and mathematics reasoning is an important facet of the body of work that describes ways in which reasoning mathematically in physics contexts is not the same as reasoning mathematically in purely mathematical contexts [130, 84, 67, 79, 116, 163, 109, 151]. Several researchers consider this difference through the lens of semiotics, highlighting the importance of symbols and the meaning they carry on physics [114, 4].

Symbols and Quantities

The symbolic form framework was developed to explain how successful physics students understand and construct equations [130]. In a study of calculus students solving problems in the contexts of areas and volumes, mathematics education researchers observed that students were more productive when they perceived symbols as representing *measures*, including both a value and a unit, throughout their reasoning [42]. This led to the introduction of the *measurement* symbolic form, which includes both a value and associated unit. The *quantity*

symbolic form, introduced in physics education research, modifies this form by including sign as an essential element of a physics quantity [160]. The quantity symbolic form represents an important difference between *quantity* in physics and *measurement* in mathematics by including connection to the physical system itself.

In mathematics education research, literal symbols (letters, sometimes loosely referred to as variables) are classified into several categories including: variables, which represent values that can vary (e.g. x, y); constants, which represent values that are always the same (e.g. π); parameters, which represent values that are not changing in that context (e.g. m and b in $y = mx + b$); generalized numbers, which are used in definitions of theorems (e.g. a and b in $a + b = b + a$); unknowns, which represent a quantity to be found (e.g. $5x = 13$); and labels, which represent units of measure (e.g. "m" for "meters") [108]. Research has demonstrated that literal symbols may invoke student difficulties due to the various roles that they play in a symbolic expression [108, 144]. In mathematics textbooks, problem solving with measures typically involves a measure symbolized by a letter, and often does not include units. The symbols are used throughout the problem; the units are declared at the outset, and tagged on to the solution of the problem. Units are not typically central to mathematical reasoning as part of instruction.

Physics, in contrast, often assumes that symbols carry physical information [116]. For example, positive and negative signs play an essential role in physics symbolizing in a way that is sometimes distinct from mathematics, and take on different meanings depending on the physical context [164]. Physics symbolizing can be challenging for novices to decode. Prior research demonstrates that physics students are more successful at solving physics problems when physical quantities are given as their numeric values instead of as literal symbols [143].

5.3.2 Framework of Covariational Reasoning in Physics

In this section, we describe the CoRP framework, shown in Table 5.2. The CoRP framework characterizes the use of covariational reasoning in physics modeling.

PROCEPTUAL UNDERSTANDING (PU)	
I. Mathematical Foundations	**II. Mathematization in Physics**
A. Mathematical Symbols	A. Mathematical Symbols
B. Common Operations	B. Constructing Quantities
C. 7 Parent Functions	1. Mathematical Structure
	2. Physical Attributes
	3. Composite Physical Quantities
	C. Variable Quantities
PHYSICS MENTAL ACTIONS (PMA)	EXPERT BEHAVIORS (EB)
PMA 1 Related Quantities	**I. Reasoning Devices**
PMA 2 Trend of Change	A. Proxy Quantity
PMA 3 Coordination of Values	B. Regions of Consistent Behavior
PMA 4 Discrete Change	C. Physically Significant Points
PMA 5 Small Chunks of Change	D. Neighborhood Analysis
	E. Compiled Models
	II. Modeling Modes
	A. Function Knowing
	B. Function Choosing
	C. Graphical Generation

Table 5.2: Current version of a framework to describe the use of covariational reasoning in physics modeling.

We divide the framework into three sections: Proceptual Understanding **(PU)**, Physics Mental Actions **(PMA)**, and Expert Behaviors **(EB)**. The Proceptual Understanding section encompasses mathematization and the interplay of quantities and models (symbolic and graphical) in both mathematics and physics contexts. The Physics Mental Actions are parallel to the Mental Actions presented by mathematics education research (see section IIA and Tab. 5.1) and are attentive to the ways in which recent work suggests that physics experts may reason differently about change and rates of change than mathematicians. Finally, the Expert Behavior section describes patterns in the ways physics experts may develop and make sense of mathematical models by using a combination of Proceptual Understanding and Physics Mental Actions; in short, the Expert Behaviors are a collection of observations of what experts do when generating models (see Chapter 3). We present these behaviors as

emergent from the evidence, suggesting how they might appear in broader contexts.

Though the framework shown in Table 5.2 is presented as having these three distinct parts, we stress that there is significant interaction between the three parts when physics experts reason about covarying quantities.

Proceptual Understanding

The Proceptual Understanding portion of the CoRP framework identifies aspects of proceptual understanding of mathematical foundations and mathematization as used in physics covariational reasoning. This section of the framework is divided into math and physics sections to attend to the foundational mathematical reasoning that students bring to physics courses, and the distinct ways that physics uses that mathematical reasoning. Attending to this distinction is an important part of physics instruction [163].

PU I. Mathematical Foundations

The Mathematical Foundations portion of the CoRP framework describes the aspects of mathematical reasoning that are necessary for productive covariational reasoning in physics. Generally speaking, introductory physics students see and learn these ideas from prerequisite math courses, though they may not display the same fluency as physics experts.

A. Mathematical Symbols is characterized by using symbols to represent mathematical concepts and *measures* (i.e., a value and an associated unit, as described by the measurement symbolic form [42]). This framework element includes, but is not limited to, symbols that represent values (e.g., x, π) and operations (e.g., $+$, $-$) as used in mathematics.

B. Common Operations is characterized by the use of operations that are ubiquitous in introductory physics and calculus classrooms, including but not limited to addition, multiplication, subtraction, division, taking a derivative or limit, and integration. Making sense of the meaning and contextual relevance of these operations is essential for covariational reasoning in physics.

C. Parent Functions is characterized by reasoning about a handful of common functions,

chosen as they represent a subset of the most common functions in introductory physics:

- linear ($y \propto x$),
- quadratic ($y \propto x^2$),
- sine / cosine ($y \propto \sin x$),
- inverse ($y \propto 1/x$),
- inverse square ($y \propto 1/x^2$),
- exponential ($y \propto e^x$), and
- logarithmic ($y \propto \ln(x/x_0)$).

We adopt the mathematics language of "parent functions," commonly used when teaching functional transformations, to illustrate that we are referring to the functional relationship between the variables. Essential background knowledge involves a general familiarity with the behavior of parent functions, including:

- the ability to sketch a graph of a given parent function, or associate a graph with a parent function,
- describe the general behavior of the function, including concavity and end behavior.
- familiarity with transforming functions (stretching, translation, etc.) in order to use them in a wide variety of scenarios.

PU II. Mathematization in Physics

Mathematization in Physics describes the foundational quantitative reasoning about physics quantities that is necessary for productive physics covariational reasoning. Introductory physics students may not have ample experience with mathematization from prior coursework.

A. Mathematical Symbols is characterized by symbolizing values and physical constants (e.g., G, π), operations and physical quantities which might be a variable, parameter, or general variable in a particular context. This also includes recognition, based on context, of a symbol as a representation of a given quantity.

B. Constructing Quantities is characterized by using common operations to construct a quantity, or to make sense of how a quantity is constructed. Here, we focus on three aspects of constructing quantities.

1. Reasoning about the *mathematical structure* of a quantity includes recognition of features of the representation of the quantity—for example, whether the quantity includes a direction or a sign.

2. Reasoning about the *physical attributes* of a quantity includes recognition of whether the quantity can be positive or negative, or discrete or continuous. The physical attributes of a quantity also includes knowledge of appropriate scale for a given context: a car will not move at an appreciable fraction of c, and a small metal sphere will not have a charge of 5 Coulombs.

3. Combining two or more quantities to create a new, *composite physical quantity* is ubiquitous in physics. Most physics quantities are product or ratio combinations of the seven base quantities (length, time, electric current, temperature, luminous intensity, mass). Graphical features such as slope (a ratio) and area under the curve (an accumulated, multiplicative quantity) are important composite physical quantities.

C. Variable Quantities is characterized by a recognition of which quantities in an equation make sense to vary, how they vary, and which other quantities don't vary. Physical models typically involve many symbols in which the literal symbols might represent constants, general variables, parameters, or varying quantities; the classification of a particular literal symbol can change from context to context—sometimes even within the same problem. De-

termining which quantities vary and which don't in a given context requires a proceptual understanding of the context.

It is common disciplinary knowledge that some quantities are more likely to change than others. For example, when characterizing the force of attraction between two massive objects ($F = \frac{Gm_1m_2}{r^2}$) it is unlikely that either mass will be manipulated as an independent variable, and more common to consider the force for different separations of the masses. A proceptual view of variable quantities includes being able to reason about which quantities are varying with respect to one another, and which represent parameters or constants. An essential part of covariational reasoning is synthesizing the relationship between two varying quantities; this requires both an understanding of the function itself (parent functions) and which quantities are relate through changes (variable quantities). Variable quantities also includes paying attention to the units of a function. For example, since a graph is a representation of how the dependent variable changes relative to changes in the independent variable, the points on the curve take on the units of the dependent variable, and the slope is a rate of change, which has units of the dependent variable over the independent variable.

Physics Mental Actions

Physics mental actions (PMA) are some identified ways that expert physicists think when directly covarying quantities: explicitly considering the change in one quantity as the result of the change in another quantity. Though the PMA are similar to the mental actions described by mathematics education researchers (see Tab. 5.1), they differ in a key way: a focus on *quantities* as the objects of covariation. This results in not only a superficial change (i.e., using the word "quantity" rather than the word "variable" in the descriptions) but also modifications that make them more consistent with the ways that physics experts use them with physics quantities. Expert-like reasoning about a quantitative relationship between quantities is generally not separable from reasoning about the physics quantities themselves [38, 168]. The physics mental actions are often guided by what is physically reasonable (see Chapter 3). For example, physics experts often rely on understanding of how a quantity can

change (e.g., continuously or discontinuously), or whether a change in one quantity would, in the real world, cause a change in another. The PMA are not hierarchical in the sense that PMA 5 is more important or more desirable than PMA 1; however, with the exception of PMA 3, the PMA are listed in order of increasing specificity about how quantities are related. PMA 3 is not included in this hierarchy because it does not include reasoning about change; we include it in the framework nevertheless because it describes a way that we relate quantities in physics.

In the covariational reasoning mental actions frameworks developed by mathematics education researchers, only one mental action (MA 4) is associated with consideration of discrete change [28, 141, 76]. All instances of discrete covariation are associated with MA 4, regardless of the size of the discrete "chunk." Different considerations of how quantities change in discrete chunks led to development of two PMA related to discrete covariational reasoning [45]. PMA 4 is most similar to what physics education research has termed scaling [8, 146, 15], and typically involves large, often integer-valued chunks. Prior work suggests that physics experts rarely if ever consider smooth, continuous changes of multiple quantities simultaneously (see Chapter 3), as described by mathematics education researchers' MA 5; instead, physics experts engaging in PMA 5 are likely to consider small "chunks" of change (e.g., considering dx to be a very small Δx) [169, 85]. For this reason, we do not include a PMA that is analogous to MA 5, but think of PMA 5 as the "most continuous" instance of discrete covariation. Consideration of small but finite changes is a common heuristic in physics modeling [87].

PMA 1: Related Quantities

PMA 1 is characterized by a recognition that one quantity is related to another quantity. Some related behaviors include labeling axes of a graph, and a verbal acknowledgement that if one quantity changes the other will as well. This typically includes a choice of which quantity is the independent quantity and which is the dependent quantity, often guided by understanding a cause and effect relationship between the quantities. The recognition that the potential energy of a spring changes as a result stretching the spring from its equilibrium

length is an example of PMA 1.

PMA 2: Trend of Change

PMA 2 is characterized by describing whether a quantity will increase or decrease as a result of another quantity increasing or decreasing. Some related behaviors include drawing arrows to indicate increases or decreases, drawing graphs that represent linear approximations, and verbalizing the trend of change. The recognition that the potential energy of a spring increases as the spring is stretched is an example of PMA 2.

PMA 3: Coordination of Values

PMA 3 involves tracking the values of two quantities to create a discrete set of associated pairs. Related behaviors involve plotting points or creating a table of values. This does not necessarily entail consideration of simultaneous change of both quantities, but rather considering multiple values of one quantity and determining the associated values of another quantity for those values. Determining the value of the potential energy of a spring for a discrete set of values of the amount that a spring is stretched from its equilibrium length is an example of PMA 3.

PMA 4: Discrete Change

PMA 4 is characterized by what is sometimes called "scaling" in physics education research. It refers to reasoning around what happens to one quantity if another, related quantity changes by a substantial, fixed amount. Some related behaviors include plugging in numbers and comparing the change, considering how the dependent quantity changes with a substantial change in the dependent quantity, and verbalizing multiplicative changes. Recognizing that doubling the stretch of a spring results in the spring potential energy increasing by a factor of four is an example of PMA 4.

PMA 5: Small Chunks of Change

PMA 5 is characterized by reasoning about the resulting change in one quantity due to small, discrete changes made to another, related quantity. This reasoning is grounded in examining what happens for *small* pieces of change. Some related behaviors include "zooming in" to a graph by examining the slope for a small region [45], verbalizing an

awareness that a change is small compared to the scale of the problem, and moving fluidly between representations of discrete change and derivative notation [85, 43]. Recognizing that the potential energy of a spring changes slowly near the equilibrium point, and more quickly further from the equilibrium point is an example of PMA 5.

Expert Behaviors

In this section, we describe a number of specific physics covariational reasoning behaviors that we call *Expert Behaviors*. These behaviors were observed in a wide range of experts completing covariational reasoning tasks described in Chapter 3 [169]. The Expert Behaviors portion of the CoRP framework is not independent of either Proceptual Understanding or the Physics Mental Actions; rather, expert behaviors are dependent on a proceptual understanding of the underlying mathematics and the relevant physics quantities, and use of the physics mental actions.

The expert behaviors described below are based on reasoning and approaches seen in physics experts' reasoning while completing graphing tasks designed to elicit covariational reasoning (see Chapter 3). As such, these behaviors can be used to construct graphical representations. We believe, however, that many of these behaviors extend naturally to the construction of symbolic representations. When possible, we describe possible extensions of the behaviors to symbolic representations of physics contexts. We note, however, that these extensions were not empirically determined.

EB I. Reasoning Devices

A "reasoning device" is a tool or small piece of reasoning that is employed while reasoning covariationally (see Chapter 3). Some of the reasoning devices are associated with the physics mental actions. In this section, we define each of the CoRP framework reasoning devices. Then we give an example of how the reasoning devices could be used together to create a graphical representation using a context typical of an introductory physics end-of-chapter problem.

A. Proxy Quantity describes when a quantity is substituted for another while covarying

two quantities. Use of a proxy quantity may allow for easier covariational reasoning. Use of proxy quantities is fundamentally related to related quantities (PMA 1), the recognition that two quantities are related.

A proxy quantity may be used to make a novel physics context into a more familiar context. For example, for constant motion contexts, time is commonly used as a proxy quantity when the task specifies "total distance traveled" as one of the quantities to be covaried (see Chapter 3). Considering how a quantity changes with respect to time is oft-practiced in physics; that is, time is a very frequently the independent quantity involved when related quantities (PMA 1) is engaged. Using time as a proxy quantity in this way often allows for easier, more rapid covariational reasoning, or the application of a familiar model. Use of time as a proxy quantity is consistent with work by mathematics and chemistry education researchers [73, 121].

A proxy quantity may also be used because of how particular quantities are understood. For example, electric potential is typically understood as a function of relative position r (i.e., $V = V(r)$). When asked to covary potential V with a quantity other than r, it is more familiar to use r as a proxy quantity (see Chapter 3). This use of proxy quantity differs from the use of time as a proxy quantity in one important way: r is not necessarily directly proportional to the quantity that it substitutes. This, again, engages PMA 1, as some quantities are considered to be functions of another quantity, and are therefore necessarily related.

B. Region of Consistent Behavior describes when a section of a given physical context is identified where the behavior of a system is constant or consistent in some way. An expert may break up a physics context into regions of consistent behavior in order to associate a single trend, function, or model with each section as appropriate. An expert may consider each region separately when constructing a graphical or symbolic representation. Sometimes, regions are determined by whether a quantity is increasing, decreasing, or constant; in these cases, this behavior is related to trend of change (PMA 2). For example, a graphical representation of a car speeding up steadily, then moving at a steady speed, and finally slowing

down steadily between two stoplights would be represented by three distinct regions on a groah of v vs t.

C. Physically Significant Point describes when a point that holds physical meaning for a given context is identified. An expert may choose and plot a small number of physically significant points to begin to construct a model. Examples of physically significant points are: a boundary between regions of consistent behavior, a bound of a quantity (a local or absolute maximum or minimum), or a point where a quantity changes the most or least rapidly. These points can be used to guide the construction of either a graphical or symbolic representation. Identification of physically significant points is an example of coordination of values (PMA 3).

D. Neighborhood Analysis involves direct covariation of quantities around physically significant points. This device is typically used in the construction of a graphical representation by considering the rate of change of the quantities around chosen points, and drawing a small line segment centered on the point to indicate the slope of the graph at that point. In other cases, neighborhood analysis can be used to check the appropriateness of a constructed representation. Experts may consider whether a constructed graphical representation models the behavior of the quantities correctly near physically significant points. This reasoning could be extended to the construction and checking of a symbolic representation by considering how the derivative of an expression should behavior at certain values; for example, whether the derivative should be positive, negative, or zero at a point. Neighborhood analysis involves reasoning about small chunks of change (PMA 5), as it involves direct covariation of quantities for a small deviations near a physically significant point.

E. Compiled Models are usually-rapid, almost-automatic associations between a relevant parent function and a given physics context. Experience with physics problem-solving often includes having an almost-automatic association between a physics context and a function (see Chapter 3). For example, a physics expert may associate: circular motion with a sinusoidal function; the trajectory of a projectile with a quadratic function; or the electric potential near a point charge with the function $1/r$. While use of compiled model does require

a trivial use of reasoning about related quantities (PMA 1), we consider direct covariational reasoning to be absent from use of a compiled model.

Example Application: To show how the reasoning devices can be used to create a graphical representation, we use the example shown in Fig. 5.2. The task asks for a graph of the gravitational potential energy of the car-Earth system for a toy car on a track

While this specific question has not been used as an interview task with physics experts, the tasks in the related study are quite similar in structure (see Chapter 3). Thus, we are not claiming that the following is evidence of expert reasoning. We present what follows to exemplify use of the reasoning devices (EB IA–EB IE) as was observed with similar tasks (see Chapter 3).

To begin this task, one could recognize that the potential energy of the system is directly proportional to the height of the car. Therefore, the height of the car can be used as a proxy quantity (EB IA) for gravitational potential energy. Next, one could break the motion of the task into several regions of consistent behavior (EB IB): the flat portion at the top of the hill, the hill portion, the flat portion of the track between the hill and loop, and the circular loop. This could be done in tandem with the identification and plotting of physically significant points (EB IC): the potential energy is at a maximum at the beginning of the hill, when the car has traveled zero distance; the potential energy is at a minimum at the end of the hill, when the car is about half-way through its journey. Similarly, one might note the beginning of loop, top of loop, and end of loop (see Fig. 5.2).

One might then reason about what happens between the plotted points. During the hill portion, one may use neighborhood analysis (EB ID) to identify that the potential energy is changing most rapidly toward the middle of the hill, and indicate that with a steep, negative line segment about halfway between the points at the top and bottom of the hill. For the circular section, one may recognize that, assuming the cart moves with constant speed, the height is related to time (proxy quantity, EB IA), and apply a compiled model (EB IE) in which the height varies sinusoidally with time for circular motion contexts. Finally, an expert may recognize that the height changes smoothly, and connect the points with a smooth curve.

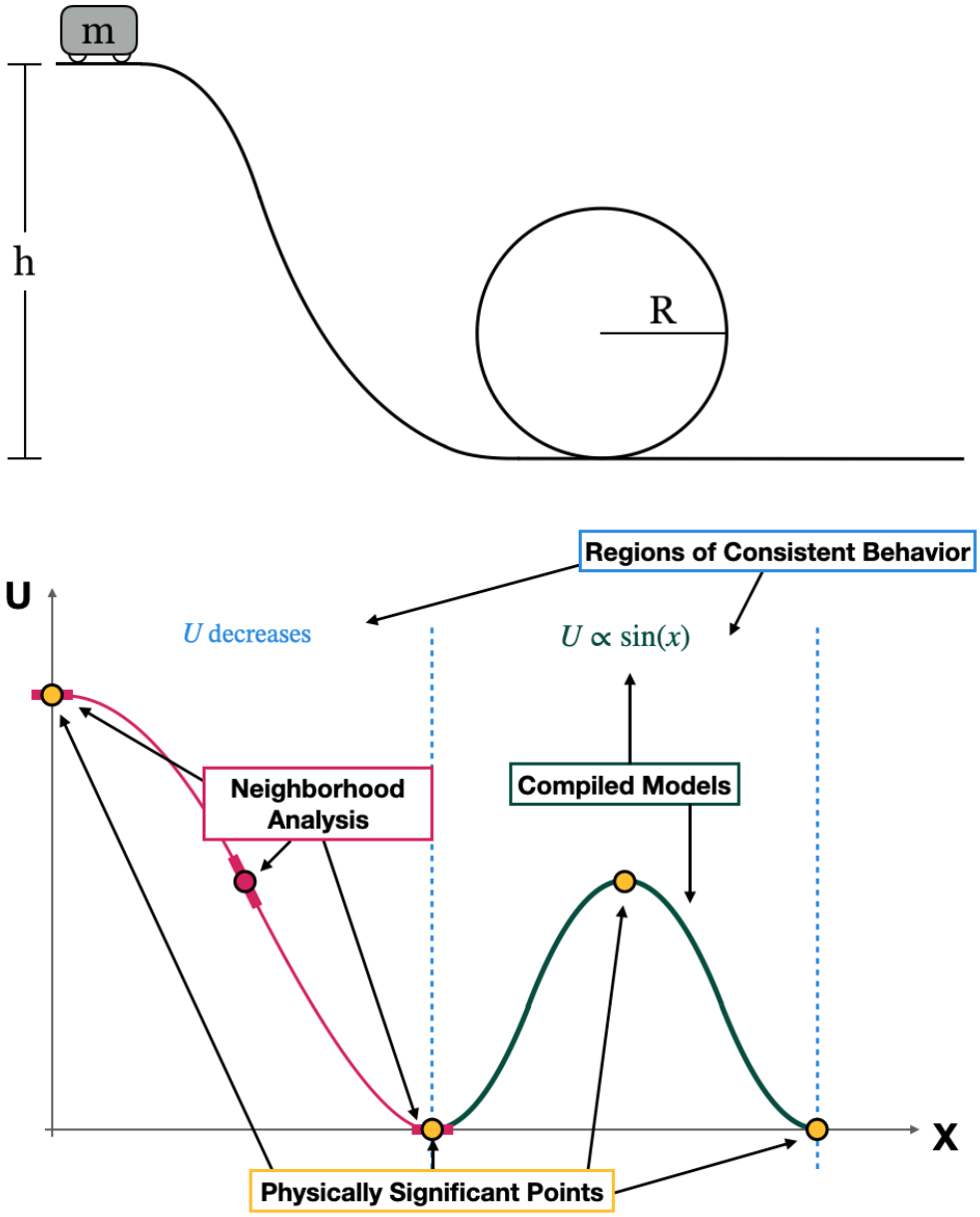


Figure 5.2: Top: Task provided here to exemplify use of the Expert Behavior reasoning devices. For this task, create a graph of the potential energy of the car-Earth system as a function of distance traveled for a toy car of mass m. The loop portion of the track is estimated as circular. Bottom: Annotated qualitatively correct graph for task.

EB II. Modeling Modes

In addition to the reasoning devices described above, we include approaches to modeling novel scenarios that have been observed in physics experts engaged in covariational reasoning tasks, referred to as *modeling modes* (see Chapter 3). Modeling modes describe the ways that experts may construct a specific mathematical representation for a given physics context. For the purposes of this work, we define a mathematical representation as a symbolic expression, or a graph, with or without an accompanying symbolic expression[1]. The modeling modes involve proceptual understanding of the relevant mathematical foundations and physics quantities, physics mental actions, and reasoning devices. In the descriptions below, we use some examples that are familiar to the reader to help illustrate the reasoning modes.

A. Function Knowing is the behavior of that relies on PMA 1 (the recognition that two given quantities are related to each other), use of a compiled model (EB IE) that associates a parent function with a given context, and proceptual understanding of the relevant parent function (PU IC) itself to associate a quantitative representation to a given physics context.

As described above, use of a compiled model involves accessing a known model for a particular physical context that carries with it a common functional relationship between two or more quantities (see Chapter 3). Function Knowing can be used to create a graphical representation by considering relevant parameters, plotting a few points based on those parameters, and connecting those points with the known function (e.g., for an object in circular motion at a steady speed, the initial conditions such as starting position and radius of the circle invoke a model for uniform circular motion). Relevant parameters can be similarly used to create a symbolic representation for the specific context. This type of "knowing" is typically guided by the physics content and draws on well-tested models of nature that are familiar to the expert.

Function knowing can be accessed in a wide range of physics contexts, based on well-

[1]We do not attempt to describe construction of covariational diagrammatic representations with this framework, though we recognize their importance.

established connections between the context and a function-based model, such as the connection between uniform circular motion and sine or cosine functions; potential near a point source and $1/r$; as well as many, many contexts with direct proportionality.

B. Function Choosing is the behavior of using a combination of physically significant points (EB IC), trend of change (PMA 2) and compiled models (EB IE) to select one of several possible functions that might be fruitful in a particular context. This process is informed by physics content knowledge of the context, and a proceptual understanding of parent functions.

For example, to generate a model of the relationship between the gravitational potential energy of a system of two objects and the distance separating the objects, one may initially identify a trend (PMA 2)—that is, that gravitation potential energy decreases as the distance between the objects decreases. Lacking additional information, assumptions about the context guide the choice of a compiled model (EB IE): assuming that the system consists of a relatively small object near the surface of a very massive planet may lead to the assumption that the change in potential energy is linear, whereas farther from the surface of the planet, the potential energy could be assumed to decrease as $1/r$. To create a graph, one may define symbolic values for initial and final positions and potential energies (EB IC), plot those points, and connect them with the function chosen based on any assumptions. Similarly, to create a symbolic expression, one could symbolize the relevant quantities, and then use them with the chosen compiled model (EB IE)—in this case, an association between the expressions $U = mgh$ and $U = -\frac{GMm}{r}$ with gravitational potential energy.

C. Graphical Generation refers to a behavior that uses direct covariational reasoning to create a qualitatively correct graph of covarying quantities. The graph then may (or may not) be used to identify an appropriate parent function that can be used as the basis of a symbolic expression.

Graphical generation is typically activated when there is no known quantitative model for a context (see Chapter 3). One approach begins with the identification of physically significant points (EB IC) and neighborhood analysis (EB ID), which involves direct covariation of

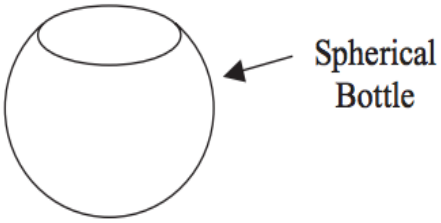


Figure 5.3: Bottle-filling task figure

quantities around the physically significant points (PMA 5). Small line segments representing the approximate rate of change of the height with respect to time around those points can be drawn and then connected together in ways deemed appropriate for the context.

Consider the *Spherical Bottle* task shown in Fig. 5.3, in which water is poured at a constant rate into a spherical bottle of radius R. Variations of the bottle-filling task have been used to assess students' covariational reasoning [28, 71]. To create a graph of how the height h of the water in the bottle varies with time, one could consider three physically significant points (EB IC): a point near the base of the bottle, a point at the middle of the bottle, and a point near the top of the bottle. The height of the water at these points will be 0, R, and $2R$, respectively. After plotting these points, neighborhood analysis (EB ID) with direct covariation of height and time (PMA 5) around them can be done . At the top and bottom, where the bottle is more narrow, the height will change quickly. At the middle, where the bottle is wider, the height will change more slowly. Based on this analysis, small line segments representing the rate of change of the height with time can be added around the identified points. The line segments should be steeper at the points at $h = 0$ and $h = 2R$ and less steep at $h = R$. Given that the height should be changing smoothly, with no discontinuities or sharp points, the line segments can be connected with a smooth curve. An expert may (or may not) identify that the created graph looks like a cubic function, and create a symbolic expression that represents the graph.

Other Tools

There are a number of other behaviors that support physics covariational reasoning that are ubiquitous in physics modeling, such as consideration of limiting cases and symmetries. While these behaviors may serve to simplify or guide covariational reasoning, they are not unique to covariational reasoning. We note that uses of these behaviors in problem solving and modeling are well-characterized by others in physics education research [158, 26, 6, 111], and we therefore focus on reasoning related to direct covariational reasoning.

5.4 Relevance of the covariational reasoning framework

In this section we demonstrate how the CoRP framework can be used to analyze student reasoning. We begin with analyses of math students' covariational reasoning, as reported in mathematics education literature. These analyses are useful because many students first develop covariational reasoning in math courses before coming to physics. As such, we suggest that algebra and calculus students are representative of students coming into introductory physics, and analysis of their covariational reasoning through the lens of the CoRP framework can help us identify the resources they bring with them from mathematics courses. These analyses also indicate opportunities for targeted instruction to develop physics students' physics covariational reasoning. We then analyze work of both introductory and upper-division undergraduate physics students from current literature, to identify the physics covariational reasoning used by more advanced students. By looking at the reasoning of students through the lens of the CoRP framework, starting with students enrolled in math courses and progressing to upper-division physics students, we can look for trends in physics covariational reasoning development with progressively more physics instruction.

5.4.1 Identifying resources of algebra and calculus students

Mathematics education researchers Johnson, McClintock, and Hornbein report on a case study of high school algebra student "Ana" [73], who participates in a number of interviews

using various covariational reasoning tasks. For a variation of a bottle-filling task (similar to the task shown in Fig. 5.3, above), Ana is shown an animation of water being poured into a spherical bottle with a cylindrical neck at the top, and is asked to create a graph to relate the amount of water in the bottle and the height of the water in the bottle. Initially, Ana creates a linear graph, saying "the height would still increase... and the volume would still increase." This is consistent with identifying a trend of change (PMA 2), as Ana is recognizing that the height increases as the volume increases. After some probing from the interviewer, Ana creates a piecewise linear graph, saying,

> ...the water is filling up very fast, and then it's like slower, and then it's really fast. Really fast would happen here [indicating neck of bottle], and kind of fast here [indicating bottom of bottle], and slow here [indicating middle of spherical portion of bottle].

Through the lens of the CoRP framework, we characterize this statement as proto-expert-like: Ana is beginning to identify regions of consistent behavior (EB IB), and is doing an early version of neighborhood analysis (EB ID) by considering differences in the rates at which the height changes in different regions in the bottle. However, rather than considering "instantaneous" rates of change around physically significant points (EB IC), Ana seems to consider average rates over larger regions. According to the sections of transcript provided by the researchers, Ana does not consider what the physical significance of a linear (or piecewise linear) graph would be and does not generally consider the meaning of the slope of the graphs she constructs.

Mathematics education researchers Carson, Jacobs, Coe, Larsen, and Hsu report on 20 college-level calculus students' performance on the same variation of the bottle-filling task (a spherical bottle with cylindrical neck) [28]. Carlson et al. report that only one student ("Student D") created a linear graph, and the student justifies their graph saying "as the volume comes up, the height would go up at a steady rate." While it is tempting to classify this reasoning as indicative of PMA 2 (Trend of Change), we believe that because the student

makes a connection between the linearity of their graph and the identified "steady" rate that the reasoning is more consistent with application of a compiled model (EB IE) or function knowing (EB IIA). While a compiled model is not associated with direct covariational reasoning, it is typically built from familiarity with and understanding of a given context and function. In this case, use of a straight line to indicate a constant rate implies understanding of why a constant slope is associated with a constant rate. We note here that a linear function is a common guess when students are asked to graph any behavior in physics.

The remainder of the students in the study by Carlson et al. created non-linear graphs, suggesting an awareness that the height of the water is not directly proportional to the amount of water in the bottle. A majority of the remaining students created graphs that were increasing and entirely concave-up or concave-down. "Student B" explains their concave-down graph by saying,

> Okay, the more water, the higher the height would be ... Right here [indicates bottom of bottle] the height will be zero and the volume is zero. As you go up, a little more height increases and the volume increases quite a bit, so the amount by which the height goes up is not as much. Once you get there [indicates halfway up the spherical part of the bottle], the height increases even slower... So, every time you have to put more and more volume in to get a greater height towards the middle of the bottle and once you get here [pointing to the top of the spherical portion], it would be linear, probably. So, it's always going up [tracing his finger along the concave-down graph], then it would be a line.

Student B begins by indicating a trend of change (PMA 2) and by coordinating values (PMA 3) at a physically significant point. The student continues by engaging in a preliminary version of neighborhood analysis (EB ID). Expert-like neighborhood analysis involves identification of physically significant points, and direct covariation of quantities in small regions around those points; here, Student B instead considers regions moving from the bottom of the bottle to the top, and compares relative rates of change in those regions. Unlike

Ana, Student B does seem to distinguish between average and instantaneous rates of change. Student B associates the shape of their graph with a perceived decreasing rate of change, having to add "more and more volume" for a given change in height. This is suggestive of reasoning about small chunks of change (PMA 5). When reasoning about the cylindrical neck of the bottle, Student B's statement that "it would be linear, probably," is consistent with a compiled model relating cylinders and volume proportional to height.

The algebra and calculus students discussed here used linear functions in their graphs. Linear functions are a resource for students coming into calculus and physics courses [172] and are often a default assumption for unfamiliar contexts for both students and experts [169, 151, 168]. We see Ana, an algebra student, activating resources representative of those that algebra students may have coming into physics classes; they can reason using the spectrum of Physics Mental Actions for linear relationships but may not yet have practiced these skills with non-uniform rates of change. While we don't consider reasoning about rates of change for linear functions to be less important than non-linear functions, we do recognize that reasoning about non-constant rates of change is more challenging for students in courses in both mathematics and physics. Facility with linear functions appears to be a strong resource to build on for students enrolled in both algebra- and calculus-based physics courses.

The calculus students studied by Carlson et al. seem to be more comfortable with changing rates of change, and show some indications of associating a linear function with a familiar physical context (i.e. the neck of the bottle). However, both groups are still inconsistent with how they make sense of the physical meaning of the quantities, rates of change, and functions they use to generate their graphs. Work in mathematics and physics education research implies that successful modeling relies on making sense of the relevant quantities as a continuous activity throughout the modeling process [37, 169, 168]. Student D and Student B create qualitatively incorrect graphs, and do not indicate an awareness that the rate of change may change in different ways for different parts of the bottle. However, both interpret the slopes of their constructed graphs as a rate of change (i.e., how the height changes with

increasing volume). This is indicative of some sensemaking about graphical representations of the physical context that could be leveraged for more expert-like neighborhood analysis and physics covariational reasoning as a whole.

5.4.2 Identifying resources of physics students

In physics, where quantities are often not easily visualized, making sense of the meaning of quantities becomes more difficult and more important to learning. In a study of 14 undergraduates enrolled in the second semester of a calculus-based introductory physics course, Bajracharya and Thompson describe how physics students reason about tasks that can be solved by finding the area under the curve in a graphical representation [11]. Many of the students recognized that the tasks could be completed by computing a (symbolic) definite integral, but no symbolic functions were provided with the graph. Students often failed to connect the integrals they constructed with the area under the curve of the graph they were given, despite the mathematical equivalence of the symbolic and graphical representations.

The first task asked students to determine a change in temperature over a given time interval from a graph of the time rate of change of temperature ($\frac{dT}{dt}$) vs. time (t). One student, "Kara," quickly produces the expression $\int \frac{dT}{dt} dt$, and is able to use dimensional analysis to check that the expression will give her a change in temperature. This is consistent with proceptual understanding of quantity in physics: Kara conceptualizes differentials as quantities with associated units (PU IIA). However, Kara fails to associate the integral with the area under the curve, and is unable to use the units associated with graphical features to guide her for this context. Kara demonstrates resources with unit analysis (PU IIA) that might be built upon to develop reasoning about graphical representations (PU IIB3).

In the second task, students were asked to find potential difference ΔV using a graph of electric field vs position. Bajracharya and Thompson report that while some students were able to state that the area under the curve would give them the quantity that the task asked for, many "were also unable to see the deeper connections between the integrals and the area under the curve" [11]. Other published research indicates that making this connection

is more difficult when the relevant quantities are more abstract and less familiar [155, 10]. Bajracharya and Thompson do not report of any of the students using consideration of units to guide their reasoning for the electrostatics task, which we view as an indication that the students' reasoning is less grounded in the meaning of the quantities.

A proceptual understanding of quantity continues to be relevant for contexts involving more advanced physics. Physics education researchers Van den Eynde, Schermerhorn, Deprez, Goedhart, Thompson, and De Cock describe a case study of second-year physics and math students reasoning covariationally in the context of the heat equation [151]. In the study, the researchers analyze the students' work through the lens of *conceptual blending*. Conceptual blending [52] describes subconscious input spaces—in this case, one physics-like and one math-like—that blend together into a single thought. Recent work has demonstrated this framework is a useful way to interpret the work of physics students and experts [13, 67, 137, 68, 151]. In the Van den Eynde et al. study, students were prompted to generate graphs of the relationship between heat flow and time based on the information in the task.

One pair of students,"Evan" and "David," appear to engage in function choosing (EB IIB):

> Evan: It [the graph] flattens out again, because that, that is just what things do under influence of the heat equation, but how would it look like at the end? Something like, eh, straight line or so maybe? I don't really know... What makes most sense?
>
> ...
>
> David: Wait, the end should be kept at 50, so there should be some kind of heat source.
>
> Evan: Yes.
>
> David: Yes, and if there is a heat source, heat should be flowing there... the beginning of the rod is so cold that it stays zero and so it cancels each other out

a bit... I think you will just get a straight line.

Yes, and if there is a heat source, heat is again flowing over there, to the beginning of the rod (6).(referring to the boundary condition at position 0 (7)) But the beginning of the rod is so cold that it stays zero (8) and so it cancels each other out a bit (9) and I think you will just get a straight line (10).

Evan: Yes, I also have the feeling it will become a straight line, because ehm, if it is not going to be a straight line, what else?

In this exchange, the students begin by using neighborhood analysis (EB ID) and making sense of the behavior of the graph at the boundaries (PMA 5). They reason about the rate of change by discussing the physical nature of the quantities. They then connect the points by choosing a function based on the trend of the quantities and the physical context (function choosing, EB IIB). The students agree that a linear relationship is likely appropriate here. Like the math students discussed above, a linear relationship between quantities is a default assumption. However, the students are continually relating their representation to the physical quantities and context: David suggests that the fixed temperatures at the ends of the rod means that there is something similar at each end so that "it cancels each other out a bit," leading to a straight line. Van den Eynde et al. suggest that many of the student statements are indicative of a well-formed conceptual blend between the mathematical machinery and relevant physics contexts. Analysis through the lens of the CoRP framework, which shows how important consideration of quantity is to function choosing, supports their findings.

Students entering physics courses are likely to have experience with covariational reasoning from pre- or co-requisite courses in mathematics. However, covariational reasoning in math courses may be focused more on coordinating changes in variables, that may or may not represent physical quantities, and less on making sense of the physical implications. Indeed, some mathematics education researchers suggest that covariational reasoning in math courses should be situated in contexts involving quantities such as height and distance that are readily visually measurable [73], so that students are able to focus on the rate of change

of a quantity without having to wrestle with the meaning of the quantity itself. While this is in line with learning objectives of math courses—and helps build a necessary foundation for physics covariational reasoning—the goals of physics courses are different. The CoRP framework was built from research that implies that a proceptual understanding of physics mathematization is also required for expert-like physics covariational reasoning.

5.5 Conclusion

Physics covariational reasoning plays a central role in expert-like quantitative modeling and is a key aspect of what it means to "think like a physicist." Because of its focus on quantitative modeling, introductory-level physics provides a unique opportunity for the development of physics covariational reasoning for a large population of students. Therefore, reliable covariational reasoning with physical quantities is a desirable student learning outcome of introductory physics courses. This is especially true because quantitative reasoning developed by instruction in physics is transferable to "real-world" contexts, and provides a foundation for scientific literacy more generally. However, it can be challenging for instructors to recognize the ways in which their own reasoning patterns differ from the focus of the prerequisite mathematics their students have taken. The differences in mathematics and physics covariational reasoning was one significant motivation for the development of the CoRP framework.

The use of covariational reasoning in physics contexts has not before been operationalized. The CoRP framework provides one operationalization. We expect that it will help instructors recognize the role (and importance) of covariational reasoning in quantitative modeling in physics, and better address its development as an important student learning outcome. In this chapter, we have described three ways that the CoRP framework can be used to achieve this goal.

First, the CoRP framework can make clear the ways in which physics covariational reasoning is distinct from the covariational reasoning described by mathematics education researchers and taught in mathematics courses. Frameworks developed by the mathematics

education research community focus largely on what we describe as "direct covariation"—direct consideration of how changes in one variable results in changes to another. While direct covariation of quantities plays an important role in physics covariational reasoning, the interconnected structure and facets of the CoRP framework attends to research that demonstrates the inherently blended nature of mathematical and physical reasoning [44, 151, 128, 168]. The foundation of physics covariational reasoning is the proceptual understanding of the underlying mathematics and physics mathematization. The physics mental actions described in the CoRP framework are similar to the mental actions described by Carlson et al. [28], but also involve quantities, rather than variables which may be free of physical context. We argue that consideration of how quantities change with respect to each other cannot happen effectively without understanding of the quantities themselves. The expert behaviors in the CoRP framework also rely on physics content knowledge, and often guide—or limit the amount of—direct covariational reasoning.

Students may come into introductory physics courses with experience with covariational reasoning in math contexts, but as shown in section 5.4, this does not guarantee facility with physics covariational reasoning. The characterization of physics covariational reasoning provided by the CoRP framework provides guidance for leveraging the experience that students have from mathematics courses. Familiarity with how covariational reasoning is used in physics, as described in the framework, can help instructors meet their students where they are, leading to more productive reasoning and quantitative modeling in physics contexts.

Second, the CoRP framework can be used to analyze student reasoning in a variety of contexts. Such analyses can be used in a number of ways; here, we discuss two such ways. First, analysis of introductory student covariational reasoning can help instructors track changes in students' covariational reasoning that occur with instruction in physics, and can inform assessment, as it explicates facets of covariational reasoning. This allows covariational reasoning to be a truly *assessable* learning objective of introductory physics courses. Also, analysis of student reasoning can aid education researchers and curriculum developers. As demonstrated in Section 5.4, analysis of students' reasoning through the

lens of the CoRP framework allows for the identification of expert-like and proto-expert-like covariational reasoning, as well as aspects of covariational reasoning in students' zones of proximal development.

Finally, the CoRP framework can guide the development of interventions. The framework identifies proceptual understanding of the foundational mathematics and physics mathematization as a basis of physics covariational reasoning. Covariational reasoning in physics is not simply doing math with physics quantities—it requires deep understanding and facility with both the mathematics and the quantities themselves. By identifying the foundations of physics covariational reasoning, the CoRP framework provides a way to determine "essential skills" [93] that can be targeted with interventions or instruction. For example, while students may come into physics courses able to produce or interpret graphs in purely mathematical contexts, they may lack facility with the physics knowledge embedded in a graph of physics quantities. Being able to interpret the meaning of a graphical feature such as a slope or an area under a curve, or identify a quantity associated with a graphical feature, may aid physics learners in understanding graphical representations.

Student difficulties with mathematics in a physics context has long been viewed as a problem of mathematical under-preparedness. While lack of adequate practice with algebraic manipulations characterizes some students' difficulties, physics has its own work to do in helping its students learn to reason mathematically in a physics context. The work described in this chapter is situated in ongoing efforts to help build physics quantitative literacy for all physics students. We have developed the CoRP framework to help support the research and instructional communities in physics gain new knowledge and develop interventions to the end of scaffolding physics students' understanding of mathematical models through their development of covariational reasoning.

Chapter 6

ASSESSING COVARIATIONAL REASONING WITH THE PHYSICS INVENTORY FOR QUANTITATIVE LITERACY (PIQL) AND THE GENERALIZED EQUATION-BASED INVENTORY FOR QUANTITY AND NEGATIVITY (GERQN)

6.1 Introduction

In this chapter, we share two short reflections about our assessment tools—the Physics Inventory for Quantitative Literacy (PIQL) and the Generalized Equation-Based Inventory for Quantity and Negativity (GERQN). Both inventories were developed in collaboration with Suzanne White Brahmia, Alexis Olsho, Trevor Smith, Andrew Boudreaux, and Philip Eaton. For a complete description of the PIQL and its development, please see our publication [166]. The PIQL is available at PhysPort for widespread use [1]; the GERQN has just completed its first year of quantitative validation at the University of Washington and will be available in the near future.

We include these two reflections in one chapter as a small collection of the early research we have conducted into interpreting the results of the PIQL and the development of the GERQN. As this work is on-going, we share here some preliminary reflections on how the inventories may be used. The first section shares an analysis of student responses to three PIQL items to illustrate how such an assessment tool can be used towards identifying student resources. The second section explains a bit about our development process for the GERQN, and lessons we have learned from our collaboration with Mieke de Cock.

6.2 Exploring student facility with "goes like" reasoning in introductory physics

6.2.1 Introduction

A perhaps unexpected byproduct of the COVID-19 pandemic is renewed clarity on how challenging it is for many to conceptualize the exponential function. This is certainly not novel; Cu Boulder Physics Education Professor Emeritus Albert Bartlett famously stated "The greatest shortcoming of the human race is our inability to understand the exponential function"[12]. This has become a public issue in the face of the coronavirus epidemic. Headlines such as "What Does Exponential Growth Mean in the Context of COVID-19?,"[27] "The Exponential Power of Now,"[119] and "Is Poor Math Literacy Making It Harder For People To Understand COVID-19 Coronavirus?"[129] have put conceptualization of function on the national stage.

It is evident that quantitative literacy—the set of skills that support the use of mathematics to describe and understand the world—is important, and lacking, in the United States today. Quantitative literacy has many facets, including reasoning about signed quantities, proportional reasoning and *covariational reasoning*—conceptualizing change in one quantity with respect to change in another quantity—and has been studied across both Mathematics and Physics Education Research [163, 28, 164]. Introductory physics, a broadly-required college course with a focus on quantifying and modeling nature, is an excellent place to address this need.

This section describes a study of students' covariational reasoning in physics contexts. It contributes to the work in mathematics education, as well as to closing a research gap in Physics Education Research (PER), where it has been shown that reasoning in physics contexts is different from reasoning in purely mathematical ones [116, 110]. We focus on one expert-like facet of physics covariational reasoning: "goes like" reasoning, which refers to the use of proportionality to illustrate relationships between two quantities (i.e. "The force *goes like* one over r squared") [172]. We present some of the ways an expert might use this reasoning, and preliminary results from two items on an assessment designed to measure

physics quantitative literacy, the Physics Inventory for Quantitative Literacy (PIQL) [166, 1]. The results of these items that suggest students have some productive resources and emergent abilities with "goes like" reasoning from prior math courses that are not yet coordinated with resources about physical quantities in some contexts. Recommendations for future work and instruction are made.

6.2.2 Quantitative Literacy in Math and Physics

Proportional reasoning—reasoning about ratio as a quantity—has been identified as critical for success in physics by physics educators and in PER. Early PER focused on identifying specific reasoning difficulties such as the tendency to use additive, rather than multiplicative, strategies and the tendency of physics students to manipulate mathematical formalism without understanding the physical meaning of the associated quantities and operations [80, 8, 9]. By the early 1980's, studies in PER had begun to systematically document and extend this body of work by using individual demonstration interviews to explore student understanding of average velocity as the ratio $\Delta x/\Delta t$ and average acceleration as the ratio $\Delta v/\Delta t$ [145, 146, 81]. More recent work has examined the relationship between basic reasoning ability, including proportional reasoning, and the learning of physics content [35].

Work on the role and challenge of proportional reasoning in physics contexts has included attention to scaling and functional reasoning [9]. We build on this body of work by integrating the language of covariational reasoning established by the Research in Undergraduate Mathematics Education (RUME) community [138, 125, 94, 28, 99, 31]. Covariation encompasses all functions that relate two or more quantities and considers multiple ways that one can think about those relationships. For example, one can consider discrete covariation (if the radius is doubled, how does the electric field change?), or continuous covariation (how does the field change smoothly as the radius is increased?) [28]. We suggest that proportional reasoning is a subset of covariational reasoning, focused specifically on linear relationships and using ratios that have meaning as a single entity (such as velocity and acceleration).

Physics educators regularly identify "thinking like a physicist" as a goal of introductory

physics. In a 2019 study of the ways in which experts use covariational reasoning to solve introductory physics problems by Zimmerman, Olsho, Boudreaux, Loverude, and White Brahmia, it was noted that physics experts use functional reasoning by employing the "$\propto$" symbol or phrases like "goes like" to illustrate relationships in statements like Area $\propto r^2$, Force goes like $1/r^2$, etc. [172] This kind of "goes like" expert thinking is used to represent a wide variety of simplified relationships between quantities, and can be viewed through the lens of conceptual blending where physics experts are subconsciously blending together their reasoning about physics concepts and their reasoning about mathematical functions [52].

In his work on proportional reasoning, Arons asserts that the capacity for scaling and functional reasoning will not necessarily develop spontaneously [9]. Indeed, the need for curricular intervention is evident from the current literature. What is less clear is what resources and emergent abilities students *do have* regarding quantitative literacy prior to physics instruction, and what educators can do to build upon these skills to develop quantitative reasoning in their students.

6.2.3 Expert "Goes Like" Reasoning

"Goes like" reasoning refers to simplifying and making sense of the covariational relationship between two changing quantities as a single function that illustrates the behavior of an evolving system. For example, consider a classic introductory physics problem about the kinematics of a ball thrown from a cliff. An expert might reason that if the ball's initial height is increased, the final speed of the ball will also increase. They might reason further that the final speed of the ball "goes like" the square root of the height. Here, "goes like" reasoning allows the expert to focus on the functional form of the relationship between two changing quantities, and to ignore any constants or pre-factors. This is a form of physics covariational reasoning as it describes how one quantity changes with respect to another in a simplified, functional way, and in turn allows for efficient problem solving, as the expected behavior of the system can be quickly and clearly illustrated.

Expert use of "goes like" reasoning is the result of a conceptual blend between their facility

with the mathematical functions involved and the physics content knowledge that enables experts to relate physical phenomena to those functions [52]. Zimmerman et al. found that physics graduate students have strong associations between certain routinely used physics quantities that allow them to make inferences about relationships between quantities in a given problem, termed "compiled relationships" (discussed previously in Chapters 3-5) [172]. This simplifies problems to those they can solve more efficiently, or to which they may already know the answer from experience [172]. Unlike novices, someone with substantial experience with physics is able to make claims such as "This problem involves a potential, which goes like $1/r$" or "This looks like scattering, so I expect it to be an exponential."

Novice physics students, in our experience, often also have useful compiled relationships that model real world contexts, and we consider these to be resources for physics learning. Many of these associations seem to evolve from prior math instruction and we see emerging evidence that students may resort to solving a physics problem by "doing math," meaning that the absence of a conceptual blend is an obstacle for effective problem solving. For example, where experts may associate circular motion with sinusoidal curves, introductory physics students may more readily associate trigonometric functions with right triangles.

This led us to wonder what resources students in an introductory physics course are using to relate two quantities, and whether they include "goes like" reasoning. In particular, we asked: do students enter introductory physics with reasoning about functional behavior already formed and ready to be applied from math courses? In addition, do their "goes like" resources improve after instruction in a physics class, where instruction typically takes the form of experts modeling their reasoning and discussing it in lecture? To answer these questions, we probed students' covariational reasoning using items from an inventory currently in development: the Physics Inventory for Quantitative Reasoning (PIQL) [131, 164].

6.2.4 Assessing "Goes Like" Reasoning

The PIQL is a reasoning inventory that measures fundamental aspects of mathematical reasoning that are ubiquitous in physics modeling. The instrument began with items targeting

proportional reasoning and reasoning about signed quantities and has since grown to include items related to covariational reasoning more broadly. During its development, the PIQL has been administered over several years in a 3-quarter calculus-based introductory physics course at a large research university in the Pacific Northwest. It is given at the start of each of the quarters (Phys 121, 122, and 123), such that it serves as a pretest for each term, and we are able to measure how students improve during the course of instruction. In this section, we report on the results of two items from 918 students responses (N=326 in 121, N=309 in 122, and N=283 in 123) from the Winter 2020 Administration (PIQL 20W), which was given in-person during 50-minute class sessions.

In this section, we focus on the results of two PIQL items aimed at assessing covariational reasoning: *Flag of Bhutan* and *Ferris Wheel*. Flag of Bhutan has been a validated item on the assessment for several years, originally taken with permission from previous work on proportional reasoning [18]. We will reframe the problem and results using the language of covariational reasoning and "goes like" thinking. Ferris Wheel is a new item on the PIQL, drawn from previous work on expert covariational reasoning in mathematics and physics [65, 172]. We will provide some analysis of the results from PIQL 20W as well as some insights on student "goes like" reasoning from recent validation interviews. Other aspects of quantification and PQL are also involved in these responses, but will not be discussed in this section.

Flag of Bhutan

In the Flag of Bhutan question, students are asked what aspects of the flag would be larger by a factor of 1.5 if the length and the width were both increased by a factor of 1.5 (see Fig. 6.1). This item was originally designed as a scaling assessment to measure student facility with both linear and non-linear relationships, as some answer choices depend linearly on length and width (such as the length of the dragon's backbone, or the distance around the edge of the flag) and the answer choice "the amount of cloth needed to make the flag" depends on length times width [18]. While scaling was considered a facet of proportional reasoning

at the time, it was understood by the researchers that scaling with non-linear functions is notably different than scaling with linear relationships. We believe that this question can be re-examined in the context of discrete covariation and "goes like" thinking.

One of the challenges of the Flag of Bhutan item as written on the PIQL is that it is a multiple-choice/multiple-response (MCMR) question. Thereby, its score is low compared to other items on the PIQL because these items are scored dichotomously for comparison with other multiple-choice/single-response items [132]. However, the nature of the item does not fully account for the significantly low number of completely correct responses (26% of all students) and this rate does not change significantly throughout the introductory sequence (25% in 121, 25% in 122, and 31% in 123) suggesting that students do not improve with instruction. The benefit of MCMR items is that the percentage of students that give partially correct answers—the student chooses *at least* one correct response and no incorrect responses—tells a more nuanced story. Answer choice (a) refers to the perimeter of the flag, (b) to the amount of cloth needed to make the flag, (c) the diagonal of the flag, and (d) the length of the dragon's backbone. 74% of all students choose (a) and a combination of (c) and/or (d), which suggests that most students have some productive resources about the functional relationship between length, width, and area that could be built upon.

By examining the partially correct responses, we can infer what kinds of resources students may be using. Only 24% of students do not choose (a), which we interpret to mean that the majority of students have facility with direct linear relationships (in this case, perimeter to length and width). In contrast, 55% of students do not choose (c) and 43% of students do not choose (d), suggesting that some students may not have yet developed the resources about more complex functional relationships, such as $\sqrt{l^2 + w^2}$, or those that do not have a known functional relationship, such as the dragon's backbone, even if the result is linear. There is, however, a correlation between choices (c) and (d): 66% of students either chose both or neither. This indicates that while the diagonal of the flag can be described by a geometric function and the backbone cannot, the majority of students are able to realize they have the same dependence on length and width. These early signs of "goes like" reasoning

Figure 6.1: The Flag of Bhutan. The prompt for this item asks students to **select all of the following quantities** that are larger by a factor of 1.5 when the length and width of the flag are both increased by a factor of 1.5: (a) The distance around the edge of the flag, (b) the amount of cloth needed to make the flag, (c) the length of the curve forming the dragon's backbone, (d) the diagonal of the flag, and (e) none of these. Students are prompted to choose all answer choices that apply. We believe the correct answers to be (a), (c) and (d).

could be developed by coordinating the demonstrated student facility of linear functions to more complex relationships with direct instruction.

Ferris Wheel

Ferris Wheel asks students to choose an equation that represents how the height of a Ferris wheel cart changes as a function of the total distance it has traveled (see Fig. 6.2). This item was developed for the PIQL to directly assess "goes like" reasoning, inspired by the 2017 Hobson and Moore study on covariational reasoning in which expert mathematicians were given an animated version of a Ferris Wheel and asked to produce a graph that relates the total distance traveled by the cart and the height of the cart [65]. Variants of the Ferris Wheel problem have been used thoroughly and are well-validated in mathematics education research, and the Hobson and Moore study was recently replicated with physics experts where it was observed that physics experts reasoned differently from the mathematicians [72, 94, 106, 33, 172]. In particular, physics experts demonstrated "goes like" reasoning and

strong compiled relationships between the circular motion presented in the animation and trigonometric functions: "the height goes like a trig function" [172].

In developing Ferris Wheel for the PIQL, we were interested to see if students also held compiled relationships, and if they were using "goes like" reasoning to solve the problem. When reformatting the item as multiple choice, we chose distractors based on common covariational relationships in physics (i.e. exponential growth). We also chose distractors based on compiled relationships students may have including the Pythagorean theorem, which students associate with triangles in interviews and in open-ended version of other PIQL items, and an expression containing the circumference, which students associate with circles.

This item appears to be considerably less challenging than Flag of Bhutan, as 58% of all students answer correctly on PIQL 20W, however it is not an MCMR question so it cannot be compared directly [131]. While the majority of students in this sample answered correctly, incorrect answer choices provide insight into what productive resources students that do not answer correctly are using. The most common incorrect choices were (d) and (a) from Figure 6.2, with an answer rate of 25% and 15% respectively across all three courses. As before, the answer rate does not change significantly during the course of instruction. These results suggest that the "circumference-like" distractor and the "Pythagorean-like" distractor are appealing to a significant portion of the student population. We interpret these answer choices as early "goes like" reasoning—these are functions that are familiar to students from prior math classes, and have been fruitful in past experiences reasoning about circles and triangles. We suggest these resources could be built upon and coordinated with reasoning about circular motion through direct instruction to develop the "goes like" reasoning experts demonstrate between circular motion and trigonometry.

Student interviews were conducted as part of the validation process for new items on the PIQL. Semi-structured, individual think-aloud interviews were conducted with six introductory physics students at another public university in the Pacific Northwest. Ferris Wheel was given as shown in Figure 6.2, and students were selected on a volunteer basis.

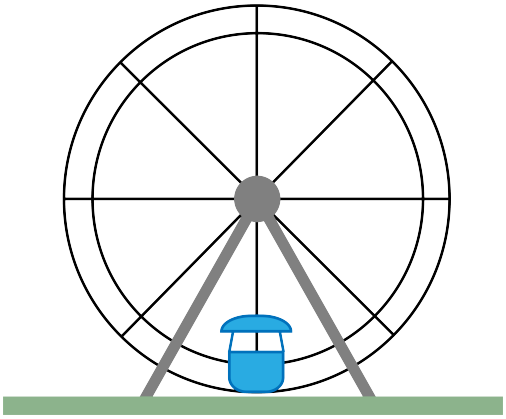


Figure 6.2: Ferris Wheel. The prompt for this item asks students to identify which expression correctly identifies how h, the height of the cart, directly changes with s, the distance traveled by the rider, where the radius of the Ferris wheel is given by R_0: (a) $h(s) = \sqrt{s^2 + R_0^2}$, (b) $h(s) = R_0 \exp(s/R_0)$, (c) $h(s) = R_0 - R_0 \cos(s/R_0)$, (d) $h(s) = s^2/(2\pi R_0)$

We do not claim that these two institutions represent identical populations; they often have slightly different average scores on PIQL assessment items, and fewer than half of those interviewed arrived at the correct answer. Therefore, the interviews provide some details into what resources students that do not yet have facility with circular motion may be using.

The Circumference-Like Distractor, answer choice (d), was highly appealing to nearly all students interviewed, citing it as familiar and associating circumference with total distance traveled: "I'd say (d) because its the only one that has $2\pi R$ in there, which is the, essentially, the circumference formula." Indeed, nearly all students interviewed began by defining the total distance traveled by the circumference, and many returned to this definition throughout their problem solving process. We recognize this as a form of quantitative reasoning—students demonstrate a strong compiled relationship between distance and circumference. They also tend to connect total distance traveled (a quantity that changes with time) with the circumference of one revolution (a quantity that is fixed), meaning that the students interviewed did not spontaneously consider the total distance *as it is changing* when using circumference to solve the problem.

The Pythagorean-like Distractor, answer choice (a), was also of significant interest to those interviewed. Every student interviewed verbally labelled this option as "Pythagorean," and many students drew an accompanying triangle, demonstrating a strong compiled relationship between the expression itself and triangular geometry. Some students used this understanding to recognize quickly that the Pythagorean approach would not work, one stating, "(A) is the Pythagorean theorem, but that doesn't make sense because that's linear distance." We interpret this as the student recognizing that the Pythagorean theorem uses linear distances, and the total distance traveled is not linear. Another student debated about the correctness of (a), stating, "This is like the Pythagorean theorem... if we do it like this, [the student draws a triangle with the hypotenuse representing total distance] I guess you could estimate [the total distance] as being a straight line." Here, we infer the student is using a method often taught in calculus courses to approximate curves as linear to make sense of the linear quantities in a circular context. Both of these students demonstrate the use of strong sense-making about linear quantities in the context of circular motion using resources from prior mathematics instruction.

The Trigonometric Answer, choice (c), was challenging for students due to its high level of complexity. Students puzzled over how to draw the corresponding triangle, "cosine gives me s over R_0...so they're saying the radius is the hypotenuse. How can that be?" We interpret this as the student sense-making about how the mathematical expression is connected to the physical representation, early evidence of student conceptual blending. Another student was unique among those interviewed in drawing a connection between the unit circle and answer choice (c), recognizing that "θ is equal to arc length over the radius...the radius should be the hypotenuse because the radius is the one thing that is measured throughout the circle." The student makes the connection of the angle as a ratio, and uses their conceptualization of the unit circle to identify the corresponding triangle.

These patterns suggest that the students interviewed have a variety of strong sense-making resources about the circumference formula, geometric approaches to the Pythagorean theorem, and early signs of conceptual blending between physical representations and math-

ematical formula. Students may not yet coordinate these resources completely correctly with circular motion, however, both in terms of time evolution and the connection between the formalism and the physical representation. Notably, those that did answer correctly during the interviews arrived at their answer by plugging in points. In particular, students focused on physically significant points, for example, the bottom and top of the Ferris Wheel where the height is at a minimum or maximum. This kind of problem solving— choosing physically relevant points to better understand the behavior of the system—has been identified as an expert-like behavior [172]. However, students uniformly did so as a last effort, suggesting they may not see the expert-like nature of this approach.

6.2.5 Conclusions

The results of Ferris Wheel and Flag of Bhutan demonstrate that while students have difficulty with physics "goes like" reasoning, they illustrate potentially productive resources that could be used to develop physics covariational reasoning. Responses to Flag of Bhutan show that students have strong "goes like" reasoning about linear relationships, and those to Ferris Wheel demonstrate that students have strong compiled relationships regarding right triangles and the Pythagorean theorem, and circles and circumference. However, without direct instruction it is challenging for students to spontaneously coordinate complex functions like $\sqrt{l^2 + w^2}$ to linear behavior and their reasoning about trigonometry to circular motion. It is important for instructors to recognize their own expert habit of thinking in terms of a small, finite number of preferred functions in physics, while students emerge from math courses without that framework. We recommend instructors consider including explicit instruction and create instructional strategies to help students develop a recognition of our preferred functions to facilitate "goes like" thinking. As this is preliminary work, additional studies are needed to more deeply understand both student and expert covariational reasoning in physics.

6.3 Assessing physics quantitative literacy in algebra-based physics: lessons learned

6.3.1 Introduction

Quantitative literacy (QL)—the interconnected skills, attitudes, and habits of mind that support the sophisticated use of familiar mathematics to describe and understand the world [139, 102]—is a central learning objective across STEM courses. Given its ubiquitous and nuanced mathematical nature, introductory physics is well-positioned to fill an educational niche of improving QL. An inventory recently developed and published by S. White Brahmia, A. Olsho, T. Smith, A. Boudreaux, P Eaton, and C. Zimmerman—the Physics Inventory for Quantitative Literacy, or the PIQL (pronounced "pickle")—has been shown to effectively measure *physics* quantitative literacy (that is, quantitative literacy in the context of physics) for calculus-based introductory physics courses [166]. However, there remains a need for an instrument to measure PQL that is appropriate for algebra-based physics courses.

Algebra-based physics courses are required for a wide variety of STEM majors, and are much more common than calculus-based courses in high schools. By developing an algebra-based inventory, we widen the population of instructors for whom the inventory is applicable and useful. As high school physics teachers are agents of change in their classrooms and the wider STEM community, making physics quantitative literacy (PQL) assessment materials available to them includes an essential group of instructors in moving the needle on QL for all students. We are therefore in the process of developing an algebra-based version of the PIQL called the Generalized Equation-based Reasoning inventory for Quantity and Negativity, or the GERQN (pronouced "gherkin"), to address this need.

The purpose of this section is to share some early insights as we explore features of physics quantitative reasoning that are ubiquitous, regardless of the mathematical preparation of the learners. We will begin with an example of how we adjusted PIQL items to be better suited for the GERQN population, and then share two lessons we learned during the process of developing the GERQN:

(I) It is well known that reasoning about the rate of change of a quantity is challenging, and learners commonly conflate it with the quantity itself [145, 146, 138, 154]. In this section we contribute to this body of work by describing how we adjusted assessment items to be within algebra-based students' zone of proximal development in graphical contexts, and offer a reflection on what kinds of resources about rates of change we think may be accessible to this population for instructors to build on.

(II) Language we use in assessment can be filtering students for the wrong reasons [54, 21, 22, 59]. Here we describe two key lessons we learned from translating the GERQN into Flemish as part of a larger project and on on-going collaboration with the researchers at Katholieke Universiteit Leuven (KU Leuven).

6.3.2 Background

Algebraic reasoning underpins both calculus-based and algebra-based introductory physics curricula. Reasoning about proportion, sense-making around symbols and representations, and making sense of quantities are common themes in the body of research on mathematical reasoning in introductory physics [116, 148, 18, 103, 130, 151, 159, 165]. The conceptual blend of procedural resources and conceptual reasoning is foundational to mathematical sense-making in physics contexts [83, 113]. Resources are pieces of knowledge that are activated when students are engaged in making sense of a particular context [41, 60, 120]. Therefore, discriminating between which resources are within the zone of proximal development for each population is essential to the development of the GERQN. Through interviews with physics faculty and a middle school mathematics education researcher, we cyclically revised and developed items for the GERQN to probe resources that are foundational to "calculus-like" reasoning at higher levels.

The GERQN development is a part of collaborative comparison study with researchers at KU Leuven in Belgium that explores similarities in, and differences between, US and Belgian students' PQL. As a part of this study, we worked with researchers in Belgium to translate

the GERQN into Flemish and gained insights about the clarity of our question statements for both native and non-native English speakers. In this section, we aim to share some lessons learned during the GERQN development process from both a mathematical and linguistic perspective, and provide suggestions for researchers and instructors to consider when drafting assessment items, creating activities, and speaking to students.

6.3.3 Development and Validation

Our procedure for inventory and item development was conducted at a large, R1 university in the Pacific Northwest, and follows the procedure used in the development of the PIQL [166, 3]. We designed the GERQN by modifying PIQL items and developing new items as needed based on the following assumed differences in the population of test-takers: (1) lower average grade level, and (2) fewer years of prior physics and mathematics courses. These items underwent expert validation with a panel of 6 experts (physics faculty who teach in the introductory physics sequence) and with a middle school mathematics education research expert. The items underwent a cycle of revisions based on expert comments and pilot testing. We then conducted student validation in 12 individual, think-aloud interviews with students from the algebra-based introductory sequence. The items were also administered in the same courses over three quarters (N = 1808). An added dimension of this process involved a collaboration with a researcher and masters students enrolled in KU Leuven's physics teacher training program. These researchers translated the GERQN in its current version into Flemish and performed 10 individual, think-aloud interviews with Belgian algebra-based physics students in Flemish. The analysis of both sets of student interviews is still ongoing. However, early results from our initial expert validation, large-N administration, and translation into Flemish have generated knowledge about bringing expert PQL into the zone of proximal development of the students we teach and assess.

Item Development and Modification

The majority of the 18 items on the GERQN are the same as those on the PIQL. For the GERQN, we rejected PIQL items that required calculus, an understanding of vector quantities, or involved physics contexts that are not taught in the algebra sequence. These items were either modified or removed.

The items "Internal Energy" and "Money" are an example of the ways PIQL items were modified for use on the GERQN (Fig. 6.3). The PIQL version of the item "Internal Energy" asks students to choose the best model to represent how work and heat are related to an object's internal energy. The PIQL item was originally developed to probe student reasoning around symbolizing (specifically the implied meanings of "+" and "–") and use of variables. The internal energy context was validated with calculus-based students, and the most popular incorrect answer for calculus-based students is to choose A ($\Delta U = Q - W$) and F ($-\Delta U = -Q + W$). These choices are incorrect because the prompt states that Q is positive when energy flows into the system, and W is positive when positive work is done on the system. Therefore, these variables should be added to represent the total change internal energy, ΔU. Student interviews reveal this is because A and F align with definitions of the first law of thermodynamics from prior instruction in which the sign of Q and/or W are defined differently. However, work, heat, and internal energy are less likely to be familiar terms for algebra-based physics students, in particular for those in pre-college classrooms.

The version "Money" was developed for the GERQN where the context was adjusted to more familiar quantities: the relationship between the change in the amount of money in a person's wallet and how much they earn or spend on goods and services. Early administrations of the GERQN suggest that the most common incorrect answer with algebra-based students is to choose all of the answer options (A: $\Delta M = -S + G$, B: $\Delta M = S + G$, C: $\Delta M = S - G$, and D: $\Delta M = -S - G$). Student interviews suggest that choosing A, B, C, and D involves reasoning that the variables represent only positive values and that a negative value would need to be indicated by an additional symbol outside the variable. Using this

(a) The internal energy of a system can be increased by doing positive work on the system or by heating it, and it can be decreased by cooling the system or if the system does work. Which of the following equations represent(s) this relationship (U is the internal energy of the system, Q is positive when energy flows into the system, and W is positive when positive work is done on the system)? ***Choose all that apply.***

a. $\Delta U = Q - W$
b. $\Delta U = -Q + W$
c. $\Delta U = Q + W$
d. $-\Delta U = Q + W$
e. $-\Delta U = Q - W$
f. $-\Delta U = -Q + W$

(b) The amount of money in your wallet changes by ΔM when you receive or spend money. The value of ΔM is greater than zero when you receive money, and the value of ΔM is less than zero when you spend money.

Let S represent the money exchanged for services. The value of S is greater than zero if you are paid for a service, and the value of S is less than zero if you pay for a service.

Let G represent the money exchanged for goods. The value of G is greater than zero if you sell something and the value of G is less than zero when you buy something.

Which of the following equations represent(s) this relationship? ***Choose all that apply.***

a. $\Delta M = -S + G$
b. $\Delta M = S + G$
c. $\Delta M = S - G$
d. $\Delta M = -S - G$

Figure 6.3: A PIQL item that has been adjusted for the GERQN due to physics context. (A) shows the original item, and (B) shows the revised item for the GERQN.

reasoning, all four equations collectively represent the possible scenarios that could occur.

This example illustrates how we made modifications to the PIQL items to adjust context and language to be more appropriate for GERQN audiences. In the following sections we will address specific lessons learned from: (I) adjusting items that measure reasoning about rates of change, and (II) adjusting language for non-native speakers based on our experience from translating the GERQN.

Reasoning about Rates of Change

Meredith and Redish summarize the mathematical reasoning foundational to algebra-based physics [92]. However, the line between what mathematical reasoning is required for calculus-based physics and algebra-based physics is not entirely clear; calculus-based physics tends to heavily rely on algebra-like reasoning with a small dose of derivatives and integrals. The items Spherical Bottle [31] and Cylindrical Bottle illustrate how we approach distinguishing between conceptual, mathematical reasoning skills for these two populations (Fig. 6.4).

Spherical bottle requires students to reason about a non-linear rate of change between the height of the water and the volume of water in the bottle—the correct graph includes not only a curve, but a curve with a change in concavity. Student difficulties with changing rates of change have been well documented in both the mathematics and physics education communities, and reasoning about changing rates of change have been shown to be an essential resource for students entering calculus [138, 28, 130, 143, 142, 23]. Based on this research, we reason that the majority of students in an algebra-based physics class may not yet have accessible resources about continuously changing rates of change. We confirmed this assertion with a local mathematics education expert at the University of Washington, who specializes in middle school level mathematics education research. To adjust the target zone, we developed the Cylindrical Bottle for the GERQN, which asks the same question but with piecewise linear graphs. Cylindrical Bottle includes only one instance where the rate of change changes—at the bottle neck—which is intended to measure how students may be reasoning about changing rates of change but within their zone of proximal development.

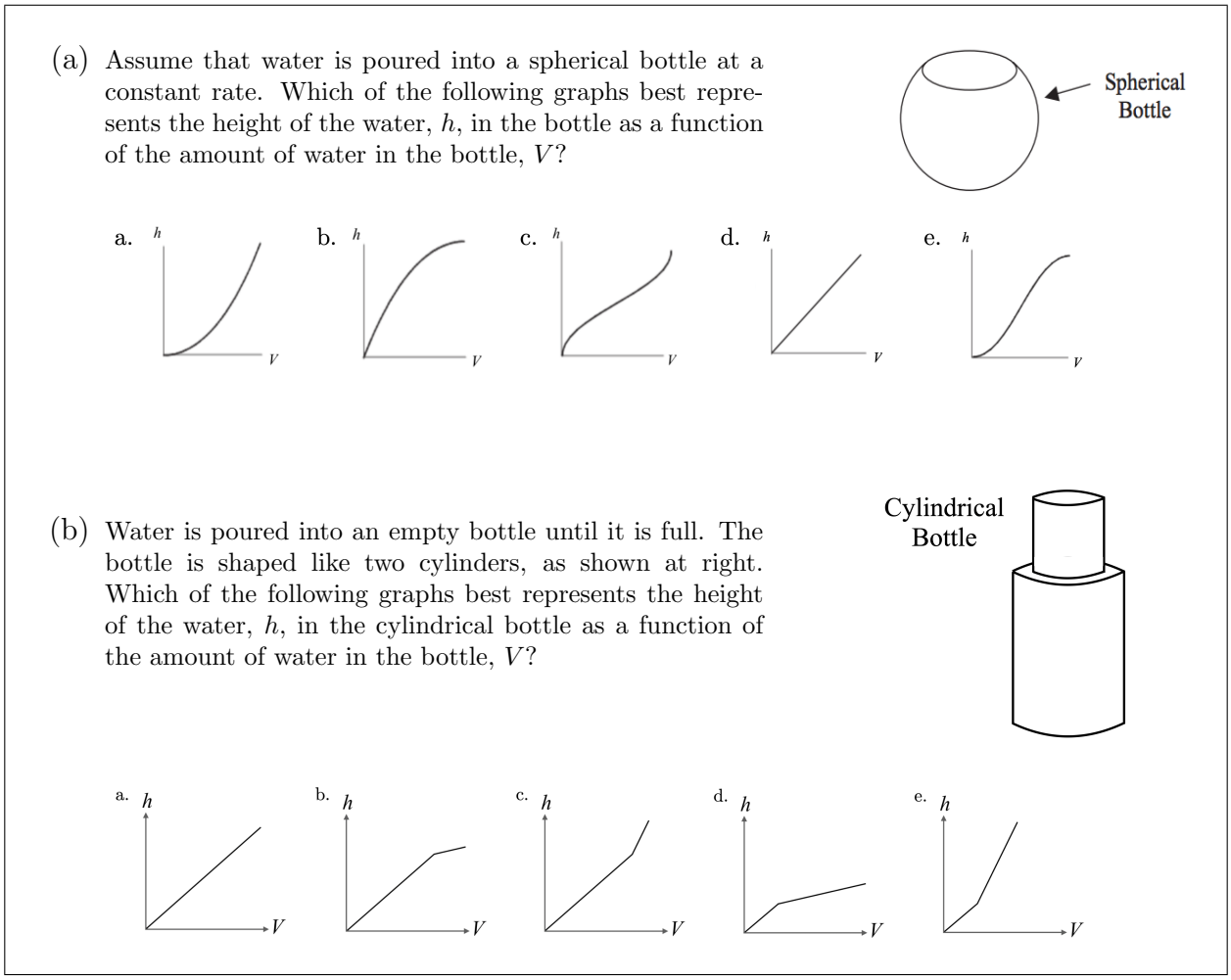


Figure 6.4: A PIQL item that has been adjusted for the GERQN due to mathematical level. (A) shows the original item, and (B) a linear version on the current iteration of the GERQN.

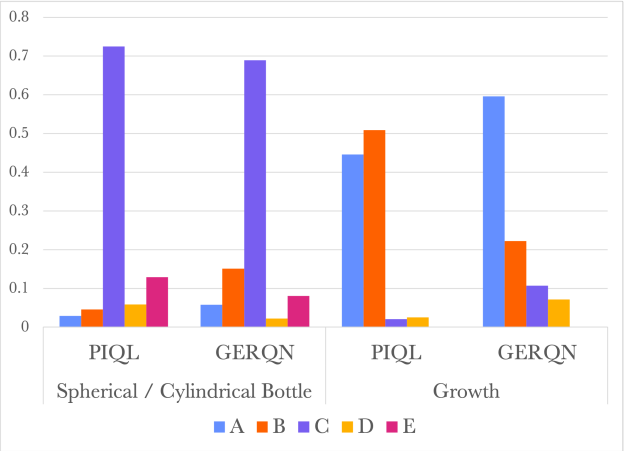
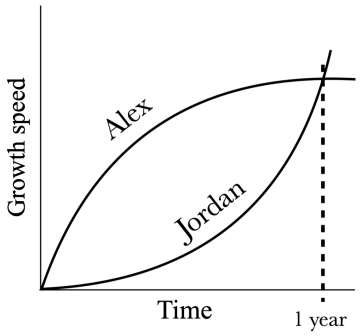


Figure 6.5: Student responses to the PIQL and GERQN versions of Spherical/Cylindrical Bottle and Growth items. The PIQL data were collected from a calculus-based introductory mechanics course (N = 240) and the GERQN data from an algebra-based introductory mechanics course (N = 225) at the same large, R1 university in the pacific northwest.

Preliminary results on the GERQN show that 68% of algebra-based students chose the correct answer on Cylindrical Bottle, compared to 72% of calculus-based students who chose the correct answer on Spherical Bottle (N = 225 for algebra-based students and N = 240 for calculus-based students, see Fig. 6.5). For calculus-based students, the most common incorrect answer is E (13% of students), where the concavity is reversed. On the GERQN version, the most common incorrect answer among algebra-based students is B (15% of students), where the rate of change in the bottle neck is slower than the rate of change in the bottle's body. While we do not suggest that we make quantitative statements by comparing across populations, the similarity in difficulty on the item and parallel structure of the most common incorrect answer suggests that the cylindrical adaptation is likely appropriate for the algebra-based sequence, and in their zone.

Some items, however, did not need to be adjusted despite appearing to some experts as though they may require calculus. The item we will refer to as "Growth", shown in Fig. 6.6, illustrates the balance between reasoning about rates of change and requiring calculus procedures to solve a problem. Growth was originally developed for the PIQL based on an

The graph at right represents how fast two children are growing vs time. The children are named Alex and Jordan, and their growth is measured starting on their 10th birthday when they are both the same height. Which of the following choices best describes how much the children have grown in the year shown?



a. Alex and Jordan have grown the same amount.

b. Alex has grown more than Jordan.

c. Jordan has grown more than Alex.

d. The graph does not provide enough information to compare how much the two children have grown.

Figure 6.6: The GERQN item "Growth," where students are asked to compare the growth of two children using a graph of growth rate. The correct answer is B, as Alex has a higher growth rate during the entire year.

item from the Precalculus Concept Assesment and prior work in PER that use the same or similarly shaped graphs, but in the context of two cars driving [31, 62, 134] . The context was modified for the GERQN and remains isomorphic to the original PIQL item. Much has been written in prior research about the intersection of these curves through both a dual process lens and other theoretical frameworks [31, 62, 134], and our data support the current understanding that the intersection is a tempting answer choice: 45% of calculus-based students and 55% of algebra-based students choose A.

However, what we find novel and most interesting about this question in our current work is a debate amongst physics experts as to whether or not the item requires calculus. We suggest that this item does not require calculus because one can reason that Alex grows at a faster rate than Jordan the entire time. However in the faculty focus group, a discussion emerged about whether this item required calculus knowledge when faculty members expressed solving the problem "required area under the curve" reasoning. Some experts

in the focus group found it challenging to reason without using calculus even though the calculus-free reasoning is more straightforward–Alex is always growing faster. After discussion between the researchers conducting the interview and the participants, unanimous agreement was reached that the item could be solved without calculus, but one faculty member added that their students would be unlikely to reason that way. We suggest that experts are so entrenched in a calculus lens of physics that for some instructors it requires more effort to reason about rates of change in any other way.

Translation into Flemish

During the translation process, we gained valuable insights into how to make the GERQN less cognitively demanding to read for non-native English speakers: (1) the nuances involved in translating specific mathematical terms such as "rate," and (2) the care required in establishing clear, consistent grammatical structures for question stems.

The first lesson we offer relates to the translation of the word "rate," which we learned from our Belgian collaborators has no direct translation in Flemish. Instead, it is translated as "speed." This made us wonder about what kinds of mathematical terms we may assume are ubiquitous across languages, and for what other languages this may be commonplace. After a quick investigation into other languages, we learned Spanish also uses the equivalent of "speed" to refer to rate of change in mathematics. For the purposes of the GERQN, this made little difference as we had already removed all instances of the word "rate" as part of lowering the required reading comprehension level. However, we include this note here as it may provide insight for other instructors as they consider writing questions and otherwise communicating with students that are non-native English speakers.

We also implemented changes across the entire inventory to ensure clear, consistent question statements. On the PIQL, question prompts vary in subtle ways, including statements such as "Select the best statement below..," "Choose the correct expression...," and "Which of the following helps figure out..." (See Table 6.1). In our discussions with Belgian researchers, it was pointed out that these varying statements unnecessarily increased the

Original	Revised
"Which one of the following best describes..." "Which of the following best represents..." "Select the choice below that best describes..." "Select the single best choice below."	"Which of the following choices best describes..."
"Which of the following expressions helps figure out..." "Which of the following is an expression for"	"Which of the following expressions can you use to find out..."

Table 6.1: Examples of ways in which question statements were rewritten on the GERQN to be more uniform and easily parsed by non-native English speakers.

workload to translate, and some words were challenging to translate exactly with the same implied meaning. In particular, many of the statements obscure the subject in question. For example, "Which of the following..." and the phrase "helps figure out" were both challenging to translate due to the nuanced grammatical structure. To address this issue, we chose to change all the question prompts to one of two uniform statements: "Which of the following choices best describes..." or "Which of the following expressions can you use to find out...", as shown in Table 6.1. These prompts were reported to be more easily understood by Belgian students, and we expect are also more easily understood by native English speaking and other non-native English speaking students alike.

6.3.4 Recommendations for Instruction

In developing the GERQN, we identified some key features of algebraic reasoning and item writing that may be useful to instructors. We interpret the results from Cylindrical Bottle to mean that graphical representations that include curvature may be out of scope for algebra students. However, piece-wise constant slope functions are in bounds, and can be used in assessment. The idea that the intersection in "Growth" (Fig. 6.6) indicates the same height

was prevalent in algebra-based students' responses as they are most likely to select "both children grow the same amount," consistent with prior research [62, 134]. We suggest the resource of recognizing sameness could be leveraged in future learning about changing rates of change, aligned with prior research [146]. We also note that an over-reliance on sophisticated mathematics can make it difficult to recognize language-based (rather than procedure-based) quantitative reasoning resources that many students can access and use to reason about rates and accumulation.

In our translation process, we identified some useful ways to consistently and clearly pose a question that was easily translatable into other languages. Employing simple grammatical structures and taking care when using specific mathematical terms such as "rate" were identified as useful first steps to ensuring language was clear for native and non-native English speakers alike.

Chapter 7

TOWARDS INTERVENTION DEVELOPMENT

In this thesis, we have presented evidence that:

- Physics quantitative literacy (PQL) is an important learning outcome of physics courses, and does not currently come along for the ride with current physics instruction.
- Physics experts reason about covariation—an essential part of PQL—differently than mathematics experts.
- Physics experts use some key behaviors when reasoning covariationally (as summarized in the CoRP) that are likely not being taught directly.
- The PIQL and GERQN are productive towards measuring student reasoning at a variety of levels to identify student resources that instruction can build on.

Through the identification of expert covariational reasoning in physics, the development of the GERQN, and the characterization of PQL throughout all levels of physics instruction, we ultimately aim to catalyze instructional change and the development and implementation of instructional materials and methods. Next steps towards this goal include:

- Finalizing the validation of the GERQN (see Chapter 7).
- Establishing validation of the PIQL and GERQN across a diverse population across socio-economic, geographical, school type, and demographic lines (see Chapter 7).
- Using the PIQL, GERQN, and CoRP together to characterize PQL, and covariational reasoning in particular, throughout all levels of physics instruction.

- Identifying appropriate learning goals based on this characterization of PQL and covariational reasoning.
- Identifying common resources amongst students enrolled in the course of interest that form the foundation of the intervention.
- Developing pilot activities to directly instruct students on this kind of reasoning with relevant physics contexts.
- Assessing the impact of these pilot activities through measuring PQL improvement via the PIQL and GERQN.
- Iterative revision until the activity appears stable and productive.

Many of these steps are on-going. In the following subsections I will provide some contextual information for how our collaboration is currently addressing, or is planning to address, each of these steps.

7.1 Finalizing the Validation of the GERQN

We follow a validation process described by physics education researchers Adams and Weiman; for a full justification and explanation of this process, please see their publication and our article describing the validatin of the PIQL [3, 166]. The validation process is mainly comprised of two parts: qualitative and quantitative validation. Qualitative validation consists of student and expert validation. In student qualitative validation, the goals are to ensure that test-takers (1) interpret the items the way you expect, (2) arrive at the correct answer using correct reasoning, and (3) do not arrive at the correct answer using incorrect reasoning. We also want to make sure we understand how students arrive at incorrect answers, so that we can interpret those results as well. To ensure qualitative validity, Adams and Weiman recommend conducting think-aloud interviews. The researcher is sure not to interrupt students' thought processes, and from their explanations we choose to either confirm that the

items are functioning as we expect, or make revisions. For expert qualitative validation, the goal is to ensure that instructors agree the test construct and the items are aligned with the goals of instruction. Important to this process is asking the instructors to both reflect on whether the items are reflective of the ways they would find valuable for their students, and to ensure that there are not ideas that are missing from the inventory that instructors consider relevant. Expert validation is conducted in focus group settings, of 4-8 experts at a time. The GERQN has undergone a full round of qualitative validation, and we expect to conduct more interviews in the fall to confirm the small number of changes that were made as a result.

The goals of quantitative validation are to ensure that the test functions the way you expect it to. This involves making sure that the test isn't too easy or too hard—if everyone gets the questions right or wrong, it doesn't provide useful information. Another goal is to make sure that the items are not redundant psychometrically, meaning that the test is as short as it can be and no shorter. We recognize that taking an inventory such as the GERQN is highly cognitively demanding, and we seek to write an assessment that minimizes this burden. Finally, we want to make sure that the assessment is measuring what we think it is measuring.

There are a series of psychometric, statistical measures for addressing these goals which depend on a large number of student responses. This spring we completed a full year administration of the GERQN, and are thus in a position to begin quantitative validity tests. Early indications are positive. For example, the difficulty (confusingly defined as % correct) of the majority of the items fall between 20% and 80% (a common, acceptable range [166]), as shown in Figure 7.1A . The discrimination (Fig. 7.1B) measures how correlated the item is to overall performance on the test. It is suggested that discrimination values be above 0.2, although it is common to relax this requirement for assessments that are intended for diverse populations [166]. On the GERQN, these early statistics appear promising—while some of our discrimination values are low, these are correlated with the items that have particularly high performance (over 80% of students get them correct). As most students are getting the

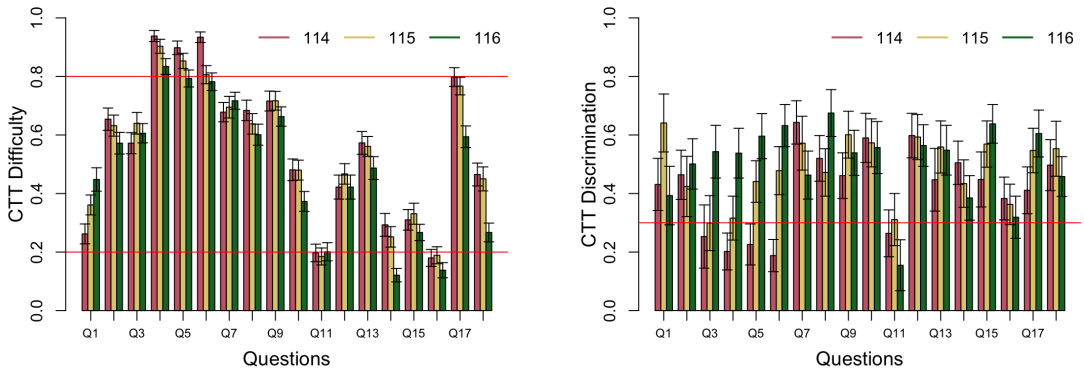


Figure 7.1: Early statistics from the 2022-2023 administration of the GERQN. These data represent a matched set of 185 students, meaning every student enrolled in the full introductory, algebra-based physics sequence. Difficulty is a measure of % correct (e.g. a difficulty of 0.8 on an item means that 80% of students chose the correct answer for that item). Discrimination is a measure of how correlated the score on that item is with the score on the overall test.

correct answer on these items, it makes sense that whether someone gets them correct is not highly correlated to their performance on the rest of the assessment. As the GERQN is being developed for all populations post-algebra I, the algebra-based sequence at the University of Washington represents some of the highest achieving students we expect to be using the assessment. Therefore, we are reserving judgement on removing items that are "too easy" until it can be tested with a more diverse pool of students.

7.2 Validity Beyond the University of Washington

Once we establish validity at the University of Washington, we will work towards quantitative validation across a wide-ranging population at other universities. We are currently engaged in partnerships with KU Leuven in Belgium, University of Georgia, North Dakota State University, Rowan University, California State University Fullerton, Everette Community College, Rowan College of South Jersey, Rowan College of Burlington County, Jackson State

University, South Seattle College, and the United States Air Force Academy. We expect data collection at these universities and colleges to begin in Fall 2023. Our statistical analysis of these data will begin as described above for quantitative validation, and we expect to also incorporate a variety of statistical measures to learn about how PQL is improving, or not, in different learning environments.

7.3 Characterizing PQL Across the Major

Simultaneous to the work described above with the GERQN, we have been administrating the PIQL to upper division courses towards being able to characterize how PQL develops. Using the results of our characterization of PQL through the physics major, we seek to better understand the specific mathematical reasoning skills in calculus-based physics that, when not learned, create barriers to students continuing the physics sequence.

Qirui Guo has led this project, and reported on early results at the Physics Education Research Conference in 2023 [**?**]. Early indications suggest that students who enter the courses with high PIQL scores are improving, and those who have relatively low PIQL scores either do not improve, or in some cases, their scores decline. This project is on-going as more data is needed to describe the patterns we are observing.

7.4 Identifying Learning Outcomes and Incoming Student Resources

Together, the PIQL (or GERQN) and the CoRP can be used to identify appropriate learning outcomes and incoming student resources. The characterization of PQL contributes to this work, and we are beginning to establish a set of goals for an introductory physics course. Our research into expert reasoning in physics also informs this process. For example, we identified that for productive learning outcomes about developing and making sense of mathematical models, explicit instruction surrounding the physics reasoning that is often implicit in mathematical modeling and about rejecting unproductive models should be considered.

We expect to use both correct and incorrect answer frequency on the PIQL and GERQN to inform our characterization. Incorrect answers on these assessments can be particularly

useful towards understanding how students reason. Our interpretation of the incorrect answers is informed by the set of validation interviews conducted previously. An example of this kind of analysis can be seen in Chapter 6. Some resources that have already been identified include reasoning about linear functions and reasoning about the pythagorean theorem. These resources could be developed into more complex sense-making about non-linear functions with appropriate instruction.

7.5 Designing Instructional Interventions

This work contributes to the development of instructional interventions in two ways: it supports the work our research group has planned for our next few years in designing activities that target PQL, and it supports current instructors to reflect on their own practices and use their experience to develop activities that work for their students. In this section, I will describe some of the future work we have planned towards activity development, but we will also be working in parallel on a dissemination plan to facilitate other instructors using this work.

Mathematics education research has already shown success developing instructional materials that develop students' conceptual mathematical reasoning in pre-calculus and calculus contexts [32, 101, 100, 140]. These materials draw heavily on research about covariational reasoning, in particular, the framework developed by Carlson et al. discussed in the background section of this thesis [28]. These materials have begun to be implemented outside of the universities in which they were developed, and early results suggest they are effective at improving covariational reasoning and quantitative reasoning in mathematics contexts [91, 66].

We expect to develop analogous activities in physics contexts to directly instruct students on this kind of reasoning with relevant physics contexts, building on Suzanne White Brahmia's prior extensive experience developing and validating mathematical reasoning materials and methods as part of our current NSF grant [50, 77]. These materials will also be informed by the results of the research described in this thesis namely addressing the blended nature

of mathematics in physics contexts, using the CoRP as a framework for learning outcomes and resources, and attending to model rejection as an essential part of developing and making sense of mathematical models. Model rejection has a prominent role in the Integrated Science Learning Environment curriuculum designed by Etkina and Heuvelen [47]. We use their work as a foundation for the development of future activities that can be more easily incorporated into already existing curricular structures. The development of pilot activities is a part of my future role as a post-doctoral scholar at the University of Washington.

In addition, as part of her post-doctoral research, Alexis Olsho in collaboration with myself, Andrew Boudreaux, Joe Olsen, and Suzanne White Brahmia has identified a series of learning goals associated with graphing and scaling based on our research into covariational reasoning in physics. These learning goals have informed the development of an online intervention using the STEM Fluency platform, which has been administered at the United States Air Force Academy. Olsho has also been using these learning goals and identified resources to develop laboratory activities that target student sense-making around uncertainty. She is currently analyzing the results of these two interventions through the lens of the PIQL, and, in this iterative process, revising and re-evaluating the goals of instruction and resources that students have. I expect to continue to work with Alexis in this collaboration towards finding effective means of developing PQL.

7.6 Assessment and Iterative Review

The materials developed will be piloted in the algebra-based and calculus-based introductory physics course at the UW, after the GERQN has been validated. We will use pre- and post-instruction scores on the GERQN and PIQL to measure the impact of the activities in improving students' PQL. An iterative practice of development, assessment, and modification will be used to develop instructional activities that are demonstrably effective at moving the needle of students' PQL.

Chapter 8

CONCLUSION

Quantitative literacy is an essential part of how people operate in a modern world. Physics, with its emphasis on modeling the real world, has the opportunity to play a key role in helping students develop their quantitative literacy. This thesis represents a collection of work that together provides a road map for developing instructional interventions that address quantitative literacy by meeting students "where they are at" as they enter introductory physics courses.

Covariational reasoning is a fundamental part of developing and validating a quantitative model. While covariational reasoning has been well-studied in mathematics contexts, it has yet to be characterized in physics. In Chapter 3, we report on empirical observations of how physics experts engage in covariational reasoning as a starting point towards characterizing physics covariational reasoning. We found that physics experts were more likely to apply already known models to a novel situation in order to seek the most reasonable solution. If no model was available, physics experts were likely to reason about the rate of change at a few, key points and connect those points with a smooth line. Throughout all of their reasoning—both while applying a familiar model and while generating a new one—physics experts were reasoning mathematically while simultaneously grounding their reasoning in the meaning of the quantities.

In Chapter 4, we report on this blended nature of physics reasoning: the physics experts reasoned mathematically with consistent reference to the physical world. This result is supported by prior work in mathematics education research [37, 128] and is aligned with a conceptual blending framework [52]. The study provides experimental evidence of the importance of quantification to physics reasoning and of the inseparability of mathematics

and physics reasoning. These findings inform how we consider covariational reasoning in physics—it includes not only particular behaviors associated with physics experts, such as applying already known models, but must also take into account the physical meaning of the quantities that are related.

Chapter 5 combines the results of Chapters 3 and 4 with prior work in physics and mathematics education research to characterize and operationalize covariational reasoning into the Covariational Reasoning in Physics (CoRP) framework. This framework provides clarity on a topic that is often elusive to instructors and researchers—as the definition of covariation is often used broadly, the framework describes specific elements of covariational reasoning that are particularly applicable to introductory physics contexts. In this way, the framework provides a path forward to characterize student reasoning, develop learning objectives for physics courses, and design interventions that meet students where they are in order to work towards learning objectives related to covariation.

Finally, Chapter 6 is focused on assessment: in order to develop effective instruction, assessment tools are required to ensure the interventions are operating as the instructor expects. The PIQL is an assessment developed by White Brahmia et al. Chapter 6.1 describes how the PIQL can be used to determine the zone of proximal development for students towards developing instructional interventions and assessing their effect. We found that many students had strong resources around linear functions that could be developed into reasoning about other functions. Through the lens of the CoRP, this is consistent with the discussion of prior research described in Sec. 5.4: the algebra and calculus students had strong resources with linear functions and large "chunks" of change that can be used as a jumping off point towards facility with non-linear functions and instantaneous rates of change.

However, the PIQL requires some mathematical reasoning that is more appropriate in calculus-based physics classrooms. As a significant fraction of physics students are enrolled in algebra-based classrooms, we realized a need for an assessment tailored to those levels. The GERQN was developed as an algebra-based version of the PIQL, and is currently under

validation at the University of Washington. In Chapter 7.2, we describe the early development process of the GERQN and emerging results that may indicate future research into language use to accommodate broader populations in our assessment tools. Together, the PIQL and the GERQN provide a set of tools that can be used across a wide variety of students towards helping instructors and researchers evaluate the quality of their instruction with respect to developing quantitative reasoning.

Including direct instruction on covariational reasoning in physics will help students broadly in their quantitative reasoning. Covariational reasoning is a blended way of reasoning, and therefore making physical sense of mathematical models is an essential part of reasoning covariationally in physics. However, it is clear from prior research that these skills are not well developed in many physics classrooms, even those that use research-based curricula [166]. As mathematical reasoning in physics contexts is distinct from mathematical reasoning in purely mathematics contexts, this kind of reasoning needs to be taught directly in physics courses.

There is emerging evidence that physics majors are eager for improved instruction that helps them understand the physical meaning of the mathematics they perform [**?**]. Developing instructional materials about covariational reasoning, grounded in the physical meaning of the quantities, would directly target this need. In addition, many students are required to take physics courses for other STEM majors because reasoning quantitatively and graphically are central learning objectives. Striving to help students improve their covariational reasoning by including it in introductory physics curriculum provides students with the tools they want and need in a wide variety of STEM fields, and in their lives beyond academia.

An opportunity emerges to help students feel supported and encouraged in a physics course when we consider including instruction on quantitative reasoning directly. Introductory physics is often perceived as difficult, and many students are likely to feel as though they do not belong. This feeling is amplified when assumptions are made about what students are familiar with coming into a course. Attending explicitly to the kinds of quantitative reasoning that we know leads to physics expertise in regular instruction helps include every

student in a physics classroom, mitigating feelings that they may not belong. This may lead to including a broader population of physics learners, and possibly a broader participation in the field.

BIBLIOGRAPHY

[1] Physics Inventory of Quantitative Literacy (PIQL).

[2] Otter, 2019.

[3] Wendy K. Adams and Carl E. Wieman. Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, 2011.

[4] John Airey and Cedric Linder. Social semiotics in university physics education. In *Multiple representations in physics education*, pages 95–122. Springer, 2017.

[5] Ayo Harriet Akatugba and John Wallace. Mathematical Dimensions of Students' Use of Proportional Reasoning In High School Physics. *School Science and Mathematics Journal*, 99(1):31–41, 1999.

[6] A. Akinyemi, M. Loverude, and J. Thompson. Solution evaluation strategies used by first-year physics students. *Physics Review Physics Education Research, in review.*

[7] Jonas Bergman Ärlebäck. On the use of realistic Fermi problems for introducing mathematical modeling in school. *The Mathematics Enthusiast*, 6(3):331–364, 7 2009.

[8] A. B. Arons. Cultivating the capacity for formal reasoning: Objectives and procedures in an introductory physical science course. *American Journal of Physics*, 44(9):834–838, 9 1976.

[9] Arnold B Arons. *A Guide to Introductory Physics Teaching.* Wiley, New York, NY, 1990.

[10] Rabindra R. Bajracharya, Vicki L. Sealey, and John R. Thompson. Student Understanding of the Sign of Negative Definite Integrals in Mathematics and Physics. *International Journal of Research in Undergraduate Mathematics Education*, 4 2023.

[11] Rabindra R. Bajracharya and John R. Thompson. Analytical derivation: An epistemic game for solving mathematically based physics problems. *Physical Review Physics Education Research*, 12(1), 3 2016.

[12] Albert Bartlett. Arithmetic, Population and Energy, 1969.

[13] Thomas J. Bing and Edward F. Redish. The cognitive blending of mathematics and physics knowledge. In *Physics Education Research Conference 2006 Conference Proceedings*, volume 883, pages 26–29, 2007.

[14] Thomas J. Bing and Edward F. Redish. Epistemic complexity and the journeyman-expert transition. *Physical Review Special Topics - Physics Education Research*, 8(1), 2 2012.

[15] J J Bissell, A Ali, and B J Postle. Illustrating dimensionless scaling with Hooke's law. *Phys. Educ*, 57(023008), 2022.

[16] Werner Blum and Dominik Leiß. 5.1 - How do Students and Teachers Deal with Modelling Problems? In Christopher Haines, Peter Galbraith, Werner Blum, and Sanowar Khan, editors, *Mathematical Modelling*, pages 222–231. Woodhead Publishing, 2007.

[17] Andrew Boudreaux, Stephen E. Kanim, Alexis Olsho, Suzanne White Brahmia, Charlotte Zimmerman, and Trevor I. Smith. Toward a framework for the natures of proportional reasoning in introductory physics. In *Physics Education Research Conference Proceedings*, pages 45–50. American Association of Physics Teachers, 2020.

[18] Andrew Boudreaux, Stephen E Kanim, and Suzanne White Brahmia. Student facility with ratio and proportion: Mapping the reasoning space in introductory physics. *ArXiv E-Prints arXiv:11511.08960 [physics.ed-ph]*, 2015.

[19] Steven Boyce, Jeffrey A Grabhorn, and Cameron Byerley. Relating students' units coordinating and calculus readiness. *Mathematical Thinking and Learning*, 23(3):187–208, 2021.

[20] Eric Brewe. Modeling theory applied: Modeling Instruction in introductory physics. *American Journal of Physics*, 76(12):1155–1160, 2008.

[21] David T. Brookes. The role of language in learning physics. *Graduate School—New Brunswick\rRutgers*, Doctor of:208, 2006.

[22] David T. Brookes and Eugenia Etkina. Using conceptual metaphor and functional grammar to explore how language used in physics affects student learning. *Physical Review Special Topics - Physics Education Research*, 3(1):1–16, 2007.

[23] Cameron Byerley. Calculus students' fraction and measure schemes and implications for teaching rate of change functions conceptually. *The Journal of Mathematical Behavior*, 55:100694, 2019.

[24] Cameron Byerley and Patrick W. Thompson. Secondary mathematics teachers' meanings for measure, slope, and rate of change. *Journal of Mathematical Behavior*, 48(February):168–193, 2017.

[25] Marcos D. Caballero, Bethany R. Wilcox, Steven J. Pollock, and Leanne Doughty. Unpacking students' use of mathematics in upper-division physics: where do we go from here? *European Journal of Physics*, 36(065004), 2015.

[26] Esmeralda Campos, Eder Hernandez, Pablo Barniol, and Genaro Zavala. Analysis and comparison of students' conceptual understanding of symmetry arguments in Gauss's and Ampere's laws. *Physical Review Physics Education Research*, 19(1), 1 2023.

[27] Matthew Cappucci. What does exponential growth mean in the context of covid-19?, 3 2020.

[28] Marilyn Carlson, Sally Jacobs, Edward Coe, Sean Larsen, and Eric Hsu. Applying Covariational Reasoning While Modeling Dynamic Events: A Framework and a Study. *Journal for Research in Mathematics Education*, 33(5):352–378, 2002.

[29] Marilyn P. Carlson. A cross-sectional investigation of the development of the function concept. In Alan H. Schoenfeld, Jim Kaput, and Ed Dubinsky, editors, *CBMS Issues in Mathematics Education*, volume 7, pages 114–162. American Mathematical Society, Providence, RI, 1998.

[30] Marilyn P. Carlson, Sean Larsen, and Sally Jacobs. An Investigation of Covariational Reasoning and its Role in Learning the Concepts of Limit and Accumulation. Technical report, 2003.

[31] Marilyn P. Carlson, Bernard Madison, and Richard West. The Calculus Concept Readiness (CCR) Instrument: Assessing Student Readiness for Calculus. 2010.

[32] Marilyn P. Carlson, Michael Oehrtman, Kevin Moore, and Alan O'Bryan. *Precalculus: Pathways to Calculus: A Problem Solving Approach*. Hayden-McNeil, MacMillian Learning, 8th edition, 2020.

[33] Carlos Castillo-Garsow, Heather Lynn Johnson, and Kevin C. Moore. Chunky and smooth images of change. *For the Learning of Mathematics*, 33(3):31–37, 2013.

[34] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser. Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5(2):121–152, 1981.

[35] V. Coletta and J. Phillips. Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics*, 73(12):1172–1182, 2005.

[36] Jere Confrey and Erick Smith. Splitting, Covariation, and Their Role in the Development of Exponential Functions. *Journal for Research in Mathematics Education*, 26(1):66–86, 1995.

[37] Jennifer A. Czocher. Introducing Modeling Transition Diagrams as a Tool to Connect Mathematical Modeling to Mathematical Thinking. *Mathematical Thinking and Learning*, 18(2):77–106, 4 2016.

[38] Jennifer A Czocher. How can emphasizing mathematical modeling principles benefit students in a traditionally taught differential equations course? *The Journal of Mathematical Behavior*, 45:78–94, 2017.

[39] Jennifer A. Czocher and Hamilton L. Hardison. Attending to Quantities Through the Modelling Space. In *Mathematical Modelling Education in East and West*, pages 263–272. Springer International Publishing AG, 2021.

[40] Jennifer A. Czocher, Hamilton L. Hardison, and Sindura S. Kularajan. A bridging study analyzing mathematical model construction through a quantities-oriented lens. *Educational Studies in Mathematics*, 111:299–321, 10 2022.

[41] Andrea A. diSessa. Toward an Epistemology of Physics. *Cognition and Instruction*, 10(2-3):105–225, 1993.

[42] Allison Dorko and Natasha Speer. Calculus Students' Understanding of Area and Volume Units. *Investigations in Mathematics Learning*, 8(1):23–46, 2015.

[43] Tevian Dray, Elizabeth Gire, Mary Bridget Kustusch, Corinne A. Manogue, and David Roundy. Interpreting Derivatives. *Primus*, 29(8):830–850, 2019.

[44] Mark Eichenlaub and Edward F. Redish. Blending Physical Knowledge with Mathematical Form in Physics Problem Solving. In *Mathematics in Physics Education*, pages 127–151. Springer International Publishing, 2019.

[45] Robert Ely and Jason Samuels. "Zoom in Infinitely": Scaling-continuous Covariational Reasoning by Calculus Students. In Aaron Weinberg, Deborah Moore-Russo, Hortensia Soto, and Megan Wawro, editors, *22nd Annual Conference on Research in Undergraduate Mathematics Education*, pages 180–187, Oklahoma City, OK, 2019.

[46] Paul J. Emigh, Reese R. Siegel, Jonathan W. Alfson, and Elizabeth Gire. Student reasoning about multivariable covariation in thermodynamics. In *Physics Education Research Conference 2019 Conference Proceedings*, pages 153–158, 2019.

[47] Eugenia Etkina and Van Heuvelen. Investigative Science Learning Environment-A Science Process Approach to Learning Physics. In Edward F. Redish and P. J. Cooney, editors, *Research-Based Reform of University Physics*, volume 1. American Association of Physics Teachers, College Park, MD, 2007.

[48] Eugenia Etkina, Aaron Warren, and Michael Gentile. The Role of Models in Physics Instruction. *The Physics Teacher*, 43(1):34–39, 1 2005.

[49] Eugenia Etkina, Aaron Warren, and Michael Gentile. The Role of Models in Physics Instruction. *The Physics Teacher*, 44(1):34–39, 1 2006.

[50] Eugenia Etkina, Suzanne White Brahmia, Rob Zisk, Jim Flakker, Danielle Buggé, Scott Robinson, Chris D'Amato, and Matt Blackman. Physics Union Mathematics, 2016.

[51] Elias Euler, Elmer Rådahl, and Bor Gregorcic. Embodiment in physics learning: A social-semiotic look. *Physical Review Physics Education Research*, 15(1), 5 2019.

[52] Gilles Fauconnier and Mark Turner. Conceptual Blending, Form and Meaning. *Recherches en Communication*, 19(2003), 2003.

[53] Rita Borromeo Ferri. Theoretical and empirical differentiations of phases in the modelling process. *ZDM — Mathematics Education*, 38(2), 2006.

[54] Markus Sebastian Feser and Dietmar Höttecke. Exploring the Role of Language in Physics Teachers' Everyday Assessment Practice. *Journal of Science Teacher Education*, 32(6):686–704, 2021.

[55] Barney G. Glaser and Anselm L. Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine Publishing Co, Chicago, 1967.

[56] Eddie M. Gray and David O. Tall. Duality, Ambiguity, and Flexibility: A "Proceptual" View of Simple Arithmetic. *Journal for Research in Mathematics Education*, 25(2):116, 1994.

[57] Gilbert Greefrath, Hans Stefan Siller, Katrin Vorhölter, and Gabriele Kaiser. Mathematical modelling and discrete mathematics: opportunities for modern mathematics teaching. *ZDM - Mathematics Education*, 54(4):865–879, 8 2022.

[58] Qirui Guo, Charlotte Zimmerman, and Suzanne White Brahmia. The Journey of Quantitative Literacy Development: Insights from Physics Majors. In *Physics Education Research Conference 2023 Poster Session*, 2023.

[59] M.A.K. Halliday and J.R. Martin. *Writing Science: Literacy and Discursive Power*. University of Pittsburgh Press, 1993.

[60] David Hammer. Student resources for learning introductory physics. *American Journal of Physics*, 68(S1):S52–S59, 2000.

[61] David Hammer and Andrew Elby. Tapping Epistemological Resources for Learning Physics. *The Journal of the Learning Sciences*, 12(1):53–90, 2003.

[62] Andrew F. Heckler. *The Role of Automatic, Bottom-Up Processes: In the Ubiquitous Patterns of Incorrect Answers to Science Questions*, volume 55. 2011.

[63] David Hestenes. Modeling games in the Newtonian World. *American Journal of Physics*, 60(8):732–748, 1992.

[64] David Hestenes. Modeling methodology for physics teachers. *AIP Conference Proceedings*, 399(1):935–958, 3 1997.

[65] Natalie L F Hobson and Kevin C. Moore. Exploring Experts' Covariational Reasoning. In *20th Annual Conference on Research in Undergraduate Mathematics Education*, pages 664–672. Moore & Thompson, 2017.

[66] Jeremiah Hower, Roneet Merkin, and Leanne Wells. A Faculty Learning Community Implements Research-based Curriculum and Pedagogy to Redesign Precalculus. *PRIMUS*, 33(5):463–486, 5 2023.

[67] Dehui Hu and N. Sanjay Rebello. Using conceptual blending to describe how students use mathematical integrals in physics. *Physical Review Special Topics - Physics Education Research*, 9(2):1–15, 2013.

[68] Tra Huynh and Eleanor C. Sayre. Blending mathematical and physical negativeness. *Proceedings of International Conference of the Learning Sciences, ICLS*, 2(2018-June):957–960, 2018.

[69] Bashirah Ibrahim, Lin Ding, Andrew F. Heckler, Daniel R. White, and Ryan Badeau. How students process equations in solving quantitative synthesis problems? Role of mathematical complexity in students' mathematical performance. *Physical Review Physics Education Research*, 13(2), 10 2017.

[70] Heather L. Johnson. Together yet separate: Students' associating amounts of change in quantities involved in rate of change. *Educational Studies in Mathematics*, 89(1):89–110, 2015.

[71] Heather Lynn Johnson. Reasoning about Quantities That Change Together. *The Mathematics Teacher*, 106(9):704–708, 5 2013.

[72] Heather Lynn Johnson, Peter Hornbein, and Sumbal Azeem. Investigating Functions with a Ferris Wheel. *The Mathematics Teacher*, 110(5):344–351, 2016.

[73] Heather Lynn Johnson, Evan McClintock, and Peter Hornbein. Ferris wheels and filling bottles: a case of a student's transfer of covariational reasoning across tasks with different backgrounds and features. *ZDM - Mathematics Education*, 49(6):851–864, 11 2017.

[74] Darrick C Jones. *Cognitive resources of physics experts*. PhD thesis, Rutgers, The State University of New Jersey, New Brunswick, NJ, 2015.

[75] Steven R. Jones. Understanding the integral: Students' symbolic forms. *Journal of Mathematical Behavior*, 32(2):122–141, 2013.

[76] Steven R. Jones. Multivariation and students' multivariational reasoning. *Journal of Mathematical Behavior*, 67(June):100991, 2022.

[77] Stephen Kanim, Suzanne White Brahmia, and Andrew Boudreaux. Physics Invention Tasks.

[78] Gülseren Karagöz Akar, İsmail Özgür Zembat, Selahattin Arslan, and Patrick W. Thompson, editors. *Quantitative Reasoning in Mathematics and Science Education*, volume 21 of *Mathematics Education in the Digital Era*. Springer International Publishing, Cham, 2022.

[79] Ricardo Karam. Framing the structural role of mathematics in physics lectures: A case study on electromagnetism. *Physical Review Special Topics - Physics Education Research*, 10(1):1–23, 2014.

[80] R. Karplus and R. W. Peterson. Intellectural development beyond elementary school II: Ratio, a survey. *School Sci. Math.*, 70(9):813–820, 1970.

[81] R. Karplus, E. K. Stage, and S. Pulos. Early adolescents' proportional reasoning on 'rate' problems. *Educational Studies in Mathematics*, 13(3):219–233, 1983.

[82] Patrick B. Kohl and Noah D. Finkelstein. Effect of instructional environment on physics students' representational skills. *Physical Review Special Topics - Physics Education Research*, 2(1), 1 2006.

[83] Eric Kuo, Michael M. Hull, Andrew Elby, and Ayush Gupta. Assessing mathematical sensemaking in physics through calculation-concept crossover. *Physical Review Physics Education Research*, 16(2), 12 2020.

[84] Eric Kuo, Michael M. Hull, Ayush Gupta, and Andrew Elby. How students blend conceptual and formal mathematical reasoning in solving physics problems. *Science Education*, 97(1):32–57, 1 2013.

[85] Mary Bridget Kustusch, David Roundy, Tevian Dray, and Corinne A. Manogue. Partial derivative games in thermodynamics: A cognitive task analysis. *Physical Review Special Topics - Physics Education Research*, 10(1), 1 2014.

[86] Michael E. Loverude. Mathematization and the 'Boas course'. In *Physics Education Research Conference 2017 Conference Proceedings*, pages 15–18, 2018.

[87] Michael E. Loverude, Stephen E. Kanim, and Luanna Gomez. Curriculum design for the algebra-based course: Just change the 'd's to deltas? In *AIP Conference Proceedings*, volume 1064, pages 34–37, 2008.

[88] David P Maloney, Curtis Hieggelke, and Stephen Kanim. nTIPERs: Tasks to Help Students "Unpack" Aspects of Newtonian Mechanics. In *Physics Education Research Conference Invited Paper Series 2010*, 2010.

[89] Catherine Martin. Mathematical Modeling: Part 1 of Series. *Colorado Mathematics Teacher*, pages 26–27, 2016.

[90] Jason M. May, Lauren A. Barth-Cohen, and Adrian A. Adams. Students' productive strategies when generating graphical representations: An undergraduate laboratory case study. In *Physics Education Research Conference Proceedings*, pages 270–276. American Association of Physics Teachers, 2021.

[91] T H McNicholl, K Frank, K Hogenson, J Roat, and M P Carlson. Improving Student Success and Supporting Student Meaning-Making in Large-Lecture Precalculus Classes. *PRIMUS*, 31(7):792–810, 8 2021.

[92] Dawn C. Meredith and Edward F. Redish. Reinventing physics for life-sciences majors. *Physics Today*, 66(7):38–43, 2013.

[93] Brendon D. Mikula and Andrew F. Heckler. Framework and implementation for improving physics essential skills via computer-based practice: Vector math. *Physical Review Physics Education Research*, 13(1):1–23, 2017.

[94] Kevin C. Moore. Quantitative Reasoning and the Sine Function: The Case of Zac. *Journal for Research in Mathematics Education*, 45(1):102, 12 2013.

[95] Kevin C. Moore, Teo Paoletti, and Stacy Musgrave. Covariational reasoning and invariance among coordinate systems. *Journal of Mathematical Behavior*, 32(3):461–473, 2013.

[96] Jayson M. Nissen, Robert M. Talbot, Amreen Nasim Thompson, and Ben Van Dusen. Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, 14(1):10115, 2018.

[97] Lorelli S. Nowell, Jill M. Norris, Deborah E. White, and Nancy J. Moules. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16(1), 2017.

[98] Tor Ole B Odden. How conceptual blends support sensemaking: A case study from introductory physics. *Science Education*, 105(5):989–1012, 2021.

[99] Michael Oehrtman, Marilyn P. Carlson, and Patrick W. Thompson. Foundational reasoning abilities that promote coherence in students' understandings of function. In Marilyn Carlson and C Rasmussen, editors, *Making the connection: Research and practice in undergraduate mathematics*, pages 27–42. Mathematical Association of America, Washington, D.C., 2008.

[100] Michael Oehrtman, Marilyn P. Carlson, and Patrick W. Thompson. Making the Connection: Research and Teaching in Undergraduate Mathematics Education. In *Mathematical Association of America*, pages 27 – 41. Mathematical Association of America, 2008.

[101] Michael Oehrtman, Michael Tallman, Zachery Reed, and Jason Martin. CLEAR Calculus.

[102] Bobby Ojose. Mathematics Literacy: Are We Able To Put The Mathematics We Learn Into Everyday Use? *Journal of Mathematics Education*, 4(1):89–100, 2011.

[103] Alexis Olsho, Suzanne White Brahmia, Andrew Boudreaux, and Trevor I. Smith. The physics inventory of auantitative reasoning: assessing student reasoning about sign. In *Proceedings of the 22nd Annual Conference on Research in Undergraduate Mathematics Education*, pages 992–997, 2019.

[104] Alexis Olsho, Charlotte Zimmerman, Andrew Boudreaux, Trevor I. Smith, Philip Eaton, and Suzanne White Brahmia. Characterizing covariational reasoning in physics modeling. In *Physics Education Research Conference Proceedings*, pages 335–340. American Association of Physics Teachers, 2022.

[105] Valerie K Otero and Danielle Boyd Harlow. Getting Started in Qualitative Physics Education Research. Technical report.

[106] Teo Paoletti and Kevin C. Moore. The parametric nature of two students' covariational reasoning. *Journal of Mathematical Behavior*, 48:137–151, 2017.

[107] Teo Paoletti and Kevin C. Moore. A covariational understanding of function: Putting a horse before the cart. *For the Learning of Mathematics*, 38(3):37–43, 2018.

[108] Randolph A Philipp. The Many Uses of Algebraic Variables. Technical Report 7, 1992.

[109] Maja Planinic, Ana Susac, Lana Ivanjek, and Željka Milin Šipuš. Comparing Student Understanding of Graphs in Physics and Mathematics. *Mathematics in Physics Education*, pages 233–246, 2019.

[110] Evan B Pollock, John R Thompson, and Donald B Mountcastle. Student Understanding Of The Physics And Mathematics Of Process Variables In P-V Diagrams. In *Physics Education Research Conference 2007 Proceedings*, 2007.

[111] Argenta M. Price, Candice J. Kim, Eric W. Burkholder, Amy V. Fritz, and Carl E. Wieman. A Detailed Characterization of the Expert Problem-Solving Process in Science and Engineering: Guidance for Teaching and Assessment. *CBE—Life Sciences Education*, 20(3), 2021.

[112] Catherine Rampell. SAT Scores and Family Income, 2009.

[113] Edward Redish. Using Math in Physics: 4. Toy models. *The Physics Teacher*, 59(9):683–688, 2021.

[114] Edward Redish and A Gupta. Making meaning with math in physics, a semantic analysis. In *GIREP Conference Proceedings*, 2010.

[115] Edward F Redish. Problem solving and the use of math in physics courses. In *World View on Physics Education in 2005: Focusing on Change Conference Proceedings*, Dehli, 2005.

[116] Edward F. Redish and Eric Kuo. Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology. *Science and Education*, 24(5-6):561–590, 7 2015.

[117] Edward F. Redish and Karl A. Smith. Looking Beyond Content: Skill Development for Engineers. *Journal of Engineering Education*, 97(3):295–307, 2008.

[118] Alan J Richards. *How students combine resources to build understanding of complex topics*. PhD thesis, Rutgers, The State University of New Jersey, 2013.

[119] Siobhan Roberts. The Exponential Power of Now, 3 2020.

[120] Amy D. Robertson, Lisa M. Goodhew, Rachel E. Scherr, and Paula R.L. Heron. University student conceptual resources for understanding forces. *Physical Review Physics Education Research*, 17(1), 3 2021.

[121] Jon-Marc G Rodriguez and Marcy H Towns. Research on Students' Understanding of Michaelis-Menten Kinetics and Enzyme Inhibition: Implications for Instruction and Learning. *The Biophysicist*, 1(2):3, 4 2020.

[122] David Rosengrant, Eugenia Etkina, and Alan Van Heuvelen. An Overview of Recent Research on Multiple Representations. In *Physics Education Research Conference Proceedings*, 2006.

[123] Rosemary S Russ, Janet E Coffey, David Hammer, and Paul Hutchison. Making classroom assessment more accountable to scientific reasoning: A case for attending to mechanistic thinking. *Science Education*, 93(5):875–891, 2009.

[124] Rosemary S. Russ, Rachel E. Scherr, David Hammer, and Jamie Mikeska. Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science Education*, 92(3):499–525, 2008.

[125] Luis A. Saldanha and Patrick W Thompson. Re-thinking co-variation from a quantitative perspective: Simultaneous continuous variation. In S. B Berensah and N. Coulombe, editors, *Proceedings of the Annual Meeting of the Psychology of Mathematics Education - North America*, 1998.

[126] Benjamin Paul Schermerhorn. *Investigating Student Understanding of Vector Calculus in Upper-Division Electricity and Magnetism: Construction and Determination of Differential Element in Non-Cartesian Coordinate Systems*. PhD thesis, University of Maine, 2018.

[127] Alan Schoenfeld and Douglas J. Herrmann. Problem Perception and Knowledge Structure in Expert and Novice Mathematical Problem Solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 1982.

[128] Kaitlyn Stephens Serbin and Megan Wawro. The Inextricability of Students' Mathematical and Physical Reasoning in Quantum Mechanics Problems. *International Journal of Research in Undergraduate Mathematics Education*, 2022.

[129] Marshall Shepherd. Is Poor Math Literacy Making It Harder For People To Understand COVID-19 Coronavirus?, 3 2020.

[130] Bruce L Sherin. How Students Understand Physics Equations. *Cognition and Instruction*, 19(4):479–541, 2001.

[131] Trevor I. Smith, Philip Eaton, Suzanne White Brahmia, Alexis Olsho, Andrew Boudreaux, Chris Depalma, Victor Lasasso, Christopher Whitener, and Scott Straguzzi. Using psychometric tools as a window into students' quantitative reasoning in introductory physics. In *Physics Education Research Conference Proceedings*, pages 560–566, 2019.

[132] Trevor I. Smith, Suzanne White Brahmia, Alexis Olsho, and Andrew Boudreaux. Developing a Reasoning Inventory for Measuring Physics Quantitative Literacy. In *Proceedings of the 22nd Annual Conference on Research in Undergraduate Mathematics Education*, pages 1181–1182, 2019.

[133] Andrzej Sokolowski. Enabling covariational reasoning in Einstein's formula for photoelectric effect. *Physics Education*, 56(3), 2021.

[134] J. Caleb Speirs, Mackenzie R. Stetzer, Beth A. Lindsey, and Mila Kryjevskaia. Exploring and supporting student reasoning in physics by leveraging dual-process theories of reasoning and decision making. *Physical Review Physics Education Research*, 17(2):20137, 2021.

[135] David Tall. The Transition to Formal Thinking in Mathematics. *Mathematics Education Research Journal*, 20(2):5–24, 2008.

[136] Henry Taylor and Michael Loverude. "I forget about physics when I go to math". In *The Learning and Teaching of Calculus Across Disciplines Conference Proceedings*, 2023.

[137] Henry Taylor and Michael E. Loverude. "So it's the same equation.": A blending analysis of student reasoning with functions in kinematics. In *Physics Education Research Conference 2018 Conference Proceedings*, volume 2018, 2018.

[138] Patrick W. Thompson. Images of Rate and Operational Understanding of the Fundamental Theorem of Calculus. *Educational Studies in Mathematics*, 26(2/3):229–274, 1994.

[139] Patrick W. Thompson. Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain, and S. Belbase, editors, *New perspectives and directions for collaborative research in mathematical education*, volume 1, chapter Quantitati, pages 33–57. University of Wyoming, Laramie, WY, 2011.

[140] Patrick W. Thompson, Mark Ashbrook, and Fabio Milner. Calculus: Newton, Leibniz, and Robinson meet technology., 2019.

[141] Patrick W. Thompson and Marilyn P. Carlson. Variation, covariation, and functions: Foundational ways of mathematical thinking. *Compendium for research in mathematics education*, (November 2016):421–456, 2017.

[142] Patrick W. Thompson, Marilyn P. Carlson, Cameron Byerley, and Neil Hatfield. Schemes for thinking with magnitudes: An hypothesis about foundational reasoning abilities in algebra. *Epistemic algebra students: Emerging models of students' algebraic knowing*, (January):1–24, 2014.

[143] Eugene T. Torigoe and Gary E. Gladding. Connecting symbolic difficulties with failure in physics. *American Journal of Physics*, 2011.

[144] Marfa Trigueros and Sonia Ursini. First-year Undergraduates' Difficulties in Working with Different Uses of Variable. *CBMS Issues in Mathematics Education*, 12, 2003.

[145] David E. Trowbridge and Lillian C. McDermott. Investigation of student understanding of the concept of velocity in one dimension. *American Journal of Physics*, 48(12):1020–1028, 1980.

[146] David E. Trowbridge and Lillian C. McDermott. Investigation of student understanding of the concept of acceleration in one dimension. *American Journal of Physics*, 49(3):242–253, 1981.

[147] Jonathan Tuminaro and Edward F. Redish. Elements of a Cognitive Model of Physics Problem Solving: Epistemic Games. *Physical Review Special Topics - Physics Education Research*, 3(2), 2007.

[148] Olaf Uhden, Ricardo Karam, Maurício Pietrocola, and Gesche Pospiech. Modelling Mathematical Reasoning in Physics Education. *Science and Education*, 21(4):485–506, 2012.

[149] Zalman Usiskin. Mathematical Modeling in the School Curriculum, 9 2011.

[150] Sofie Van Den Eynde, Benjamin P. Schermerhorn, Johan Deprez, Martin Goedhart, John R. Thompson, and Mieke De Cock. Dynamic conceptual blending analysis to model student reasoning processes while integrating mathematics and physics: A case study in the context of the heat equation. *Physical Review Physics Education Research*, 16(1):10114, 2020.

[151] Sofie Van Den Eynde, Paul Van Kampen, Wim Van Dooren, and Mieke De Cock. Translating between graphs and equations: The influence of context, direction of translation, and function type. *Physical Review Physics Education Research*, 15(2), 2019.

[152] Alan Van Heuvelen and Xueli Zou. Multiple representations of work–energy processes. *American Journal of Physics*, 69(2):184–194, 2 2001.

[153] Joëlle Vlassis. Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and Instruction*, 14:469–484, 2004.

[154] Joshua Von Korff and N. Sanjay Rebello. Distinguishing between "change" and "amount" infinitesimals in first-semester calculus-based physics. *American Journal of Physics*, 82(7):695–705, 2014.

[155] Joshua Von Korff and N Sanjay Rebello. Distinguishing between "change" and "amount" infinitesimals in first-semester calculus-based physics. *American Journal of Physics*, 82(7):695–705, 2014.

[156] L. S. Vygotsky. *Development of Higher Psychological Processes*. Harvard University Press, 1978.

[157] Eric Weber and Patrick W. Thompson. Students' images of two-variable functions and their graphs. *Educational Studies in Mathematics*, 87(1):67–85, 2014.

[158] Gary White, Tiffany-Rose Sikorski, Justin Landay, and Maryam Ahmed. Limiting case analysis in an electricity and magnetism course. *Physical Review Physics Education Research*, 19(1):010125, 4 2023.

[159] Suzanne White Brahmia. *Mathematization in introductory physics*. PhD thesis, Rutgers University, 2014.

[160] Suzanne White Brahmia. Quantification and its importance to modeling in introductory physics. *European Journal of Physics*, 40(4), 2019.

[161] Suzanne White Brahmia and Andrew Boudreaux. Signed quantities: mathematics based majors struggle to make meaning. In *Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education*, 2017.

[162] Suzanne White Brahmia, Andrew Boudreaux, and Stephen E Kanim. Developing Mathematical Creativity with Physics Invention Tasks. *arXiv e-prints*, page arXiv:1602.02033, 2 2016.

[163] Suzanne White Brahmia, Andrew Boudreaux, and Stephen E. Kanim. Obstacles to Mathematization in Introductory Physics. *ArXiv E-Prints arXiv:1601.01235 [physics.ed-ph]*, 2016.

[164] Suzanne White Brahmia, Alexis Olsho, Andrew Boudreaux, Trevor I. Smith, and Charlotte Zimmerman. A conceptual blend of physics quantitative literacy reasoning inventory items. In *Proceedings of the 23rd Annual Conference on Research in Undergraduate Mathematics Education*, pages 862–867, 2020.

[165] Suzanne White Brahmia, Alexis Olsho, Trevor I. Smith, and Andrew Boudreaux. A framework for the natures of negativity in introductory physics. *Physical Review Physics Education Research*, 16(1):10120, 2019.

[166] Suzanne White Brahmia, Alexis Olsho, Trevor I. Smith, Andrew Boudreaux, Philip Eaton, and Charlotte Zimmerman. Physics Inventory of Quantitative Literacy: A tool for assessing mathematical reasoning in introductory physics. *Physical Review Physics Education Research*, 17(2), 12 2021.

[167] Bethany R. Wilcox, Marcos D. Caballero, Daniel A. Rehn, and Steven J. Pollock. Analytic framework for students' use of mathematics in upper-division physics. *Physical Review Special Topics - Physics Education Research*, 9(2), 11 2013.

[168] Charlotte Zimmerman, Alexis Olsho, Michael Loverude, and Suzanne White Brahmia. Empirical evidence for the inseparability of mathematics from physics: expert reasoning about novel graphing tasks. *Physical Review Physics Education Research, in review*, 2023.

[169] Charlotte Zimmerman, Alexis Olsho, Michael Loverude, and Suzanne White Brahmia. Expert covariational reasoning resources in physics graphing tasks. *Physical Review Physics Education Research, in review*, 2023.

[170] Charlotte Zimmerman, Alexis Olsho, Michael E. Loverude, and Suzanne White Brahmia. Identifying Covariational Reasoning Behaviors in Expert Physicists in Graphing Tasks. In *Proceedings of the 23rd Annual Conference on Research in Undergraduate Mathematics Education*, 2020.

[171] Charlotte Zimmerman, Alexis Olsho, Suzanne White Brahmia, Andrew Boudreaux, Trevor I. Smith, and Philip Eaton. Exploring student facility with "goes like" reasoning in introductory physics. In *Physics Education Research Conference 2020 Conference Proceedings*, pages 605–610, 2020.

[172] Charlotte Zimmerman, Alexis Olsho, Suzanne White Brahmia, Michael E. Loverude, Andrew Boudreaux, and Trevor I. Smith. Towards understanding and characterizing expert covariational reasoning in physics. In *Physics Education Research Conference 2019 Conference Proceedings*, pages 693–698, 2019.

Appendix A

EXPERT COVARIATIONAL REASONING CODEBOOKS

Here we present the codebooks for Study 1 and for Studies 2 and 3 for the interested reader. In the codebook for Studies 2 and 3, an effort has been made to connect the code categories for Studies 2 and 3 back to those for Study 1 to aid the reader.

Code Category	Description	Example
Mapping Math to Physics	The epistemic game of attempting to fit a particular mathematical function to a physical scenario [13].	"And then total distance starts to increase, so we're going to go up [draws a line increasing linearly]."
Considering Quantity	Identifying, defining, and representing quantities either given in the problem statement or devised in the process of problem solving.	"So distance from Tacoma's on the y axis and then total distance, its going to be on the x axis."
Approach to the Problem	Identifying the "best way" to start a task, unprompted. This includes but is not limited to breaking down the task and identifying important points.	"So there's a moment when they're doing that loop around Tacoma looks like half a circle. So that's going to be a constant distance. So I'm going to start off with that."
Trends of Change	Relating the change in one quantity to the change in another; most commonly by discussing the rate of change or slope of a graph.	"When it's halfway to the bottom, the speed should be the greatest."
Graphical Structure	Finding a larger graphical trend between two quantities; this includes linking together known rates of change with a smooth curve, and seeking graphical features such as extrema.	"So if I want it to be slow, fast, slow, I think it looks like...[draws a curve connecting tangent lines]."
Compiled Models	Using an already-known relationship between two quantities that is strongly associated with a physical context to model a particular situation. This new term is further discussed in Section 3.4.2.	"I'm just like — oh, trig function — its a circle."

Table A.1: Study 1 code categories that emerged during data analysis.

Code Category	Description	Example
Physical World	Analogous to Mapping Math to Physics from Study 1; we limit this to when interviewees specifically discuss the physical world in the context of the task.	"This point here [gesturing to the top of the Ferris wheel], that's the maximum height."
Quantity	Analogous to Considering Quantity from Study 1; additionally includes explicit reference to the use of a quantity not named in the task to make the task easier to reason about.	"Distance traveled...uh..well, [is the] same thing as time basically."
Mental Actions	The use of the mental actions as described by mathematics education research in Table 2.1.	"So you get basically some some curve like this. Slowly, slowly changing in the bottom."
Graphing	Analogous to Graphical Structure from Study 1; includes behaviors associated with drawing a graph, such as connecting points, assuming a smooth curve, and distinguishing between independent and dependent quantities.	"Okay, total distance traveled and the electric potential it measures. So, I want to do distance traveled on the x-axis... and the y will be electric potential."
Function	Analogous to Compiled Relationships from Study 1; includes both compiled relationships and the identification of a particular relationship as a particular functional form (parent function).	"So the trajectory looks like is a quadratic—It's sort of a parabola."

Table A.2: Studies 2 and 3 code categories that emerged during data analysis. These codes were developed using thematic analysis [97] such that they describe emergent patterns from the data. While an effort was made to set aside the codebook from Study 1, many similar themes arose. We connect these code categories to those of Study 1 in the descriptions for the ease of the reader.

Appendix B

MATH-PHYSICS BLEND TIMELINES

The complete timelines for all the participants in this study.

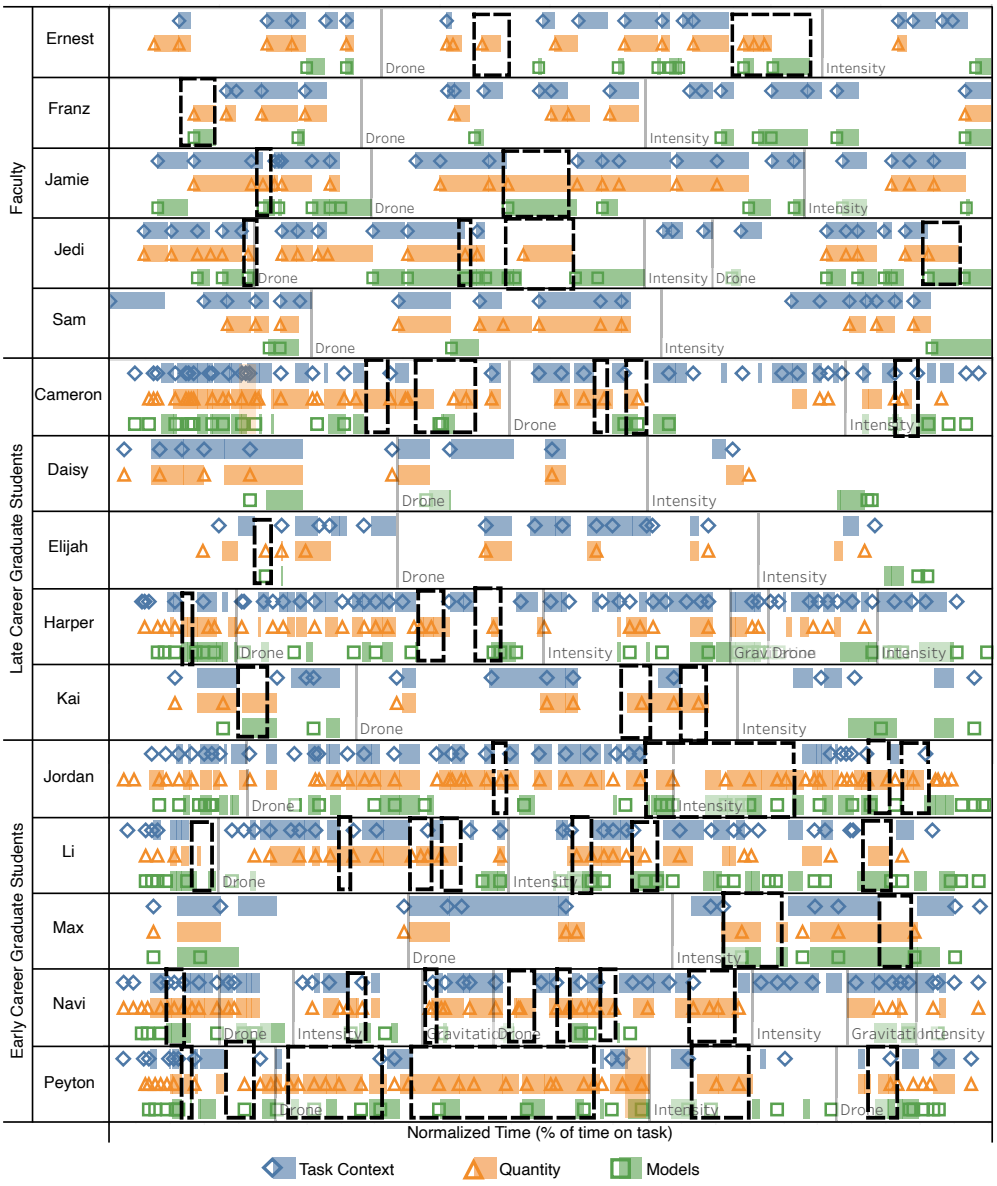


Figure B.1: A visual representation of when participant statements were assigned the "Quantity," "Task Context," and "Models" code categories. The length of the bars corresponds to the time the participant was engaged in that line of reasoning. The horizontal axis represents time, normalized to the length of each interview. Therefore, it can be thought of as a percentage of their total time on task. The outlined regions are where "Quantity" appears without "Task Context."