Development and analysis of assessments that promote sensemaking in physics

by

Amogh Sirnoorkar

M.Sc., Central University of Karnataka, India, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Abstract

Assessments are an integral part of academic environments and can present opportunities for students to make sense of novel contexts using their existing ideas. Assessments also provide insights on students' learning and the efficacy of the pedagogical practices. Consequently, physics education research shares a storied history of developing research-based assessments (RBAs) that support students' sensemaking. However, contemporary studies have noted several shortcomings of the existing RBAs such as: (i) lack of clarity for instructors in interpreting students' scores to make modifications to their instruction, (ii) misalignment between the content of the assessments and the local learning goals of the instructors, (iii) scarcity of standardized assessments for undergraduate physics, and (iv) the need to shift the focus of RBAs from "knowing" to "doing" physics. Researchers have also called for probing the contextual factors that influence students' sensemaking in physics.

In light of these observations, I present methodological and theoretical approaches to developing and analyzing next-generation RBAs. I introduce the development process of a novel RBA – the Thermal and Statistical Physics Assessment (TaSPA) – which focuses on assessing the "doing" aspect of physics along with providing actionable feedback for instructors to modify their courses. Additionally, this assessment allows instructors to choose what they wish to assess, thereby bridging the gap between assessment objectives and the local learning goals of the instructors. I elucidate the leveraging of existing theoretical and design frameworks in the development of TaSPA and how the interplay of these frameworks addresses some of the shortcomings of the contemporary assessments. This diagnostic represents a paradigm shift in how assessments are envisioned and designed by the discipline-based education research community.

I also complement the contemporary literature by theoretically exploring assessment task features that increase the likelihood of students sensemaking in physics. I identify the task features by first noting the salient characteristics of the sensemaking process as described in the science education literature. Existing theoretical ideas from cognitive psychology, education, and philosophy of science are then leveraged in unpacking the task features which elicit the characteristics of sensemaking. Furthermore, I leverage Conjecture Mapping – a framework from design-based research – to articulate how the proposed task features elicit the desired outcome of sensemaking. I argue that to promote sensemaking, tasks should cue students to unpack the underlying mechanism of a real-world phenomenon by coordinating multiple representations and by physically interpreting mathematical expressions.

The proof of concept of this idea is then presented through analysis of students' reasoning about the tasks embodying the proposed features. The analysis is then extended to provide an explicit account of the intertwining between modeling and sensemaking processes. The analyses reveal that particular aspects of modeling and sensemaking processes co-occur. For instance, the priming on the "given" information from the problem statement constituted the students' engagement with their mental models, and their attempts to resolve inconsistencies in understanding involved the use of external representations. I find that barriers experienced in modeling can inhibit students' sustained sensemaking.

Major contributions of this work include: (i) elucidating a methodological approach in developing an RBA that promotes the "doing" aspect of physics; (ii) demonstrating an agent-based perspective in exploring assessment task features; (iii) operationalizing conjecture mapping in the context of task design in physics; (iv) introducing a methodology extendable to unpack task features which can elicit other valued epistemic practices; and (v) an explicit framework-based unpacking of the association between modeling and sensemaking.

This dissertation opens up avenues for future explorations such as extending the proposed methodology in developing RBAs for other upper-division physics courses such as quantum mechanics. The presented theoretical and methodological approaches can also be extended in exploring features of the assessment tasks that promote additional epistemic practices such as argumentation and modeling. Researchers can also explore expanding the proposed list of task features, and the accompanying constraints (if any).

Development and analysis of assessments that promote sensemaking in physics

by

Amogh Sirnoorkar

M.Sc., Central University of Karnataka, India, 2016

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Approved by:

Major Professor
James T. Laverty

Copyright

Abstract

Assessments are an integral part of academic environments and can present opportunities for students to make sense of novel contexts using their existing ideas. Assessments also provide insights on students' learning and the efficacy of the pedagogical practices. Consequently, physics education research shares a storied history of developing research-based assessments (RBAs) that support students' sensemaking. However, contemporary studies have noted several shortcomings of the existing RBAs such as: (i) lack of clarity for instructors in interpreting students' scores to make modifications to their instruction, (ii) misalignment between the content of the assessments and the local learning goals of the instructors, (iii) scarcity of standardized assessments for undergraduate physics, and (iv) the need to shift the focus of RBAs from "knowing" to "doing" physics. Researchers have also called for probing the contextual factors that influence students' sensemaking in physics.

In light of these observations, I present methodological and theoretical approaches to developing and analyzing next-generation RBAs. I introduce the development process of a novel RBA – the Thermal and Statistical Physics Assessment (TaSPA) – which focuses on assessing the "doing" aspect of physics along with providing actionable feedback for instructors to modify their courses. Additionally, this assessment allows instructors to choose what they wish to assess, thereby bridging the gap between assessment objectives and the local learning goals of the instructors. I elucidate the leveraging of existing theoretical and design frameworks in the development of TaSPA and how the interplay of these frameworks addresses some of the shortcomings of the contemporary assessments. This diagnostic represents a paradigm shift in how assessments are envisioned and designed by the discipline-based education research community.

I also complement the contemporary literature by theoretically exploring assessment task features that increase the likelihood of students sensemaking in physics. I identify the task features by first noting the salient characteristics of the sensemaking process as described in the science education literature. Existing theoretical ideas from cognitive psychology, education, and philosophy of science are then leveraged in unpacking the task features which elicit the characteristics of sensemaking. Furthermore, I leverage Conjecture Mapping – a framework from design-based research – to articulate how the proposed task features elicit the desired outcome of sensemaking. I argue that to promote sensemaking, tasks should cue students to unpack the underlying mechanism of a real-world phenomenon by coordinating multiple representations and by physically interpreting mathematical expressions.

The proof of concept of this idea is then presented through analysis of students' reasoning about the tasks embodying the proposed features. The analysis is then extended to provide an explicit account of the intertwining between modeling and sensemaking processes. The analyses reveal that particular aspects of modeling and sensemaking processes co-occur. For instance, the priming on the "given" information from the problem statement constituted the students' engagement with their mental models, and their attempts to resolve inconsistencies in understanding involved the use of external representations. I find that barriers experienced in modeling can inhibit students' sustained sensemaking.

Major contributions of this work include: (i) elucidating a methodological approach in developing an RBA that promotes the "doing" aspect of physics; (ii) demonstrating an agent-based perspective in exploring assessment task features; (iii) operationalizing conjecture mapping in the context of task design in physics; (iv) introducing a methodology extendable to unpack task features which can elicit other valued epistemic practices; and (v) an explicit framework-based unpacking of the association between modeling and sensemaking.

This dissertation opens up avenues for future explorations such as extending the proposed methodology in developing RBAs for other upper-division physics courses such as quantum mechanics. The presented theoretical and methodological approaches can also be extended in exploring features of the assessment tasks that promote additional epistemic practices such as argumentation and modeling. Researchers can also explore expanding the proposed list of task features, and the accompanying constraints (if any).

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my heartfelt gratitude to my advisor, Dr. James T. Laverty (JT), for his unwavering support and encouragement throughout my academic journey over the last four years. I am grateful for his dedication in fostering an environment that was relaxed, informal, and conducive to engaging discussions on research. His attentiveness to my mental well-being and his willingness to grant me autonomy and flexibility in my research have been invaluable. I cannot overstate the importance of his positive and cheerful guidance, which has made my Ph.D. experience both successful and enjoyable.

Also thanks to my committee members, Dr. Bethany Wilcox, Dr. Jeremy Schmit and Dr. Frederick Burrack for their cooperation and constructive feedback. I am thankful for their efforts in accommodating my scheduling requests and other logistical matters.

Special thanks to Dr. Bethany Wilcox for her insights and support on the Thermal and Statistical Physics Assessment (TaSPA) project. I am grateful for her sustained, weekly participation in my research over the last three years. I would also like to take this opportunity to convey my heartfelt thanks to her for always being available to discuss and provide feedback, accommodating my requests in a generous manner.

I am also grateful to Dr. Dean Zollman for his time and efforts in going through this dissertation and providing valuable inputs. Also sincere thanks to Brandi Lohman for going through manuscripts and providing constructive feedback.

Special thanks to Kim Coy for all her support. Thanks for your unwavering and unconditional care during my graduate journey. Your handwritten greeting cards on birthdays and festivals remain testimony of your commitment towards ensuring graduate students' well-being. Any description of my graduate journey would be incomplete without you. Also thanks to Dr. Eleanor Sayre for her support in my first year of graduate study.

I am also grateful for Dr. Rosemary Russ, Dr. Qing Ryan, and Dr. Tor Odden for their insights on the theoretical constructs of epistemic framing and sensemaking.

Also thanks to current and past KSUPER colleagues for their support. Thanks to Alexander Adamson, Bill Bridges, Lydia Bender, Tyler Garcia, Dr. Hien Khong, Dr. Amali Jambuge Priyanka, Parker Poulos, and Joshua Weaver for the constructive feedback during weekly meetings. On the same note, thanks to Dr. Jessey Changstrom, Dr. Shams El-Adawy, Dr. Christopher Hass, Katelyn Hubenig, Dr. Tra Huynh, and Dr. Dina Zohrabi for their support in the last four years. Special thanks to Dr. Paul D. O. Bergeron for weekly discussions and feedback on the modeling project. Also thanks to Katherine C. Ventura for collecting interview data used in this dissertation.

I am also thankful to Dr. Katherine D. Rainey (Katie) and Michael Freeman from the University of Colorado Boulder for being excellent colleagues to work with. I will cherish weekly interactions with you on the TaSPA project.

Also thanks to Dr. Rajeev Joshi, Dr. Srikant Utagi, Sai Srikar, Dr. Raghavendra Deshpande, and Pragya Samal for their continued support.

I express my sincere thanks to faculty members of the Department of Physics at Kansas State University, Dr. Matthew J. Berg, Dr. Loren Greenman, Dr. Chii-Dong Lin, Dr. Bharat Ratra, and Dr. Uwe Thumm for their classes on the graduate courses. Their flexibility and assistance in accommodating scheduling requests were instrumental in my navigation through teaching, research, and course responsibilities. On the same note, thanks to Peter Nelson and John Pratt for providing valuable experience in teaching during my first year of graduate study. I am also grateful to the Department of Physics and the Graduate School at Kansas State University, as well as the National Science Foundation for funding the research reported in this dissertation.

I am thankful to Sai, Shweta, Biswa, Rudra, Laxmi, Gunvant, Supriya, Prakash, Snehal, Ganesh, Suyog, Deepesh. Nitin, Deepika, Priyanka, Taruna, Rajesh, Ashwin, Rakshit, Prithvi, Aabila, Ratnakshi, Gujju, and the Indian Students Association for being a family away from home. On the same note, thanks to Mehran Solemani for being an extended family member at Manhattan.

Lastly, the pillar of my support, thanks to my family without whom this dissertation would not have been possible.

Dedication

To my parents, Rekha and Srinivas Sirnoorkar. Their principles of upbringing have played a significant role in shaping my path towards completing this dissertation.

To my sister, Dr. Neeti Sirnoorkar Chagi, whose constant support has been a major driving factor in my dissertation journey.

To my beloved wife, Sneha Kulkarni, for her unconditional love and support.

Preface

This dissertation titled "Development and analysis of assessments that promote sensemaking in physics" is submitted for the degree of Doctor of Philosophy in the Department of Physics at Kansas State University. The research was conducted under the supervision of Dr. James T. Laverty. The work presented in this dissertation is original to the best of my knowledge, except where acknowledgments and references are highlighted. Most of the work has been published or has been accepted for publication or currently in review in the following peer-reviewed journals and conferences:

- **Amogh Sirnoorkar**, Paul Bergeron, and James T. Laverty, "Sensemaking and scientific modeling: Intertwined processes analyzed in the context of problem solving in physics", *Physical Review Physics Education Research*, 2023. https://link.aps.org/doi/10.1103/PhysRevPhysEducRes.19.010118.

- **Amogh Sirnoorkar**, Amali Priyanka Jambuge, Katherine D. Rainey, Alexander Adamson, Bethany R. Wilcox and James T. Laverty, "Theoretical approach for promoting scientific practices through a scalable standardized assessment for undergraduate physics", *Proceedings, International Conference on Learning Sciences (ICLS)*, 2023, (accepted).

- James T. Laverty, **Amogh Sirnoorkar**, Amali Priyanka Jambuge, Katherine D. Rainey, Joshua Weaver, Alexander Adamson, and Bethany R. Wilcox, "A New Paradigm for Research-Based Assessment Development", *Proceedings, Physics Education Conference (PERC)*, Grand Rapids, USA, 2022. https://doi.org/10.1119/perc.2022

- **Amogh Sirnoorkar**, and James T. Laverty, "A methodology for identifying task features that promote sensemaking", Proceedings, *Physics Education Research Con-*

ference (PERC), USA, 2021.
https://doi.org/10.1119/perc.2021

- Paul Bergeron, **Amogh Sirnoorkar** and James T. Laverty, “Towards an analytic framework for characterizing student use of models”, Proceedings, *Physics Education Research Conference (PERC)*, USA, 2021.
https://doi.org/10.1119/perc.2021

- **Amogh Sirnoorkar** and James T. Laverty, “Theoretical exploration of task features that facilitate student sensemaking in physics”, *Physical Review Physics Education Research*, 2023 (In review).
https://doi.org/10.48550/arXiv.2302.11478

Chapter 1

Introduction

Engaging with assessment tasks is an integral component of students' academic practice, particularly in Science, Technology, Engineering, and Mathematics (STEM) courses. Task-based learning finds its origin in situated cognition, according to which, knowledge construction is influenced by the activity, culture, and the context in which the learning occurs[5;6]. Assessments thus present opportunities for students to make sense of novel contexts in light of their existing ideas. Consequently, physics education research (PER) shares a storied history of developing and analyzing research-based assessments (RBAs) that support students' interaction with curricular content[7–9].

However, educators have noted several shortcomings of the existing RBAs such as: (i) lack of clarity for instructors in interpreting students' scores to make effective course modifications[10], (ii) a misalignment between the content of the assessments and the local learning goals of the instructors[10], (iii) the scarcity of standardized assessments for upper division physics content, particularly thermal and statistical physics[11], and (iv) the need to shift the focus of RBAs from "knowing" to "doing" physics[12–14]. Furthermore, several studies in PER have advocated for facilitating students' content understanding through "sophisticated epistemology" - the feature of leveraging different modes of reasoning while engaging with curriculum[15–19]. The community has also emphasized promoting pedagogical practices that facilitate students in generating new knowledge by connecting their existing ideas[20].

My dissertation attends to these concerns by developing and analyzing assessments that promote sensemaking in physics. Sensemaking represents the cognitive process of attempting to address a perceived gap in one's understanding[21] and assists students' in better comprehending the curricular content by leveraging different forms of knowledge and practices[22]. Sensemaking is also one of the many ways through which scientists and engineers generate new knowledge[23–25]. Given this significance, there has been an uptick in investigations on promoting and analyzing this cognitive process in classroom settings. Researchers have noted construction and critique of claims during scientific discourses[23] and vexing questions during interactions[26] to be instrumental in sustaining sensemaking. Additionally, the blending of model- and evidence-based reasoning[27] during classroom instruction, computational reasoning about physics scenarios[28], addressing quantitative problems through qualitative insights and vice-versa[29] have been observed to facilitate sensemaking in physics.

In chapter 3, I present a theoretical approach in developing a standardized assessment for upper-division physics that promotes students' sensemaking of novel contexts through application of their existing knowledge[14]. The Thermal and Statistical Physics Assessment (TaSPA) promotes the "doing" aspect of science in addition to providing actionable feedback for instructors to support their students' learning. I elucidate the approach in achieving these objectives by bringing together contemporary theoretical ideas such as Three-Dimensional Learning[13], Evidence Centered Design[30], and Self-Regulated Learning[31].

In chapter 4, I theoretically explore the features of assessment tasks that increase the likelihood of students' sensemaking in physics[32;33]. I identify the task features by first noting the salient characteristics of the sensemaking process as described in the science education literature. Existing theoretical ideas from cognitive psychology, education, and philosophy of science are then leveraged in unpacking the task features which elicit the characteristics of sensemaking. I adopt an agent-based approach in articulating the task features by shifting the vocabulary from "*tasks entailing a feature X*" to "*tasks that cue students about X"* or "*tasks that cue students to do X*". Such vocabulary would better account for students' agency along with the local practices of their learning environments. Furthermore, through a Conjecture Map[1] – a framework from design-based research – arguments about how the proposed task features elicit the desired outcome of sensemaking are elucidated. I argue that to promote sensemaking, tasks should cue students to unpack the underlying mechanism of a real-world phenomenon by coordinating multiple representations and by physically interpreting mathematical expressions. In chapter 5, I analyze students' responses to tasks embodying the above-mentioned features and present a proof-of-concept of the proposed theoretical idea.

In chapter 6, I continue analyzing students' responses to similar tasks but through an additional theoretical lens of modeling[34]. Modeling is a common cognitive process through which humans comprehend their surrounding world[35–37]. When modeling, the person engaging in modeling abstracts and simplifies the target system to facilitate an explanation or a prediction[3;4;37;38]. During this process, ideas from one's knowledge are organized and applied to reason about various phenomena. Consequently, model-based reasoning can be a crucial component of sensemaking[27]. Studies have noted that modeling and sensemaking share several common features including the objective of "to figure it out"[21]. The two processes consist of multiple phases including priming of prior knowledge of the concerned context, noticing discrepancies while reasoning, and generating new knowledge by connecting existing ideas[39;40]. However, most of these accounts focus on describing what modeling entails, but rarely articulate what constitutes as making sense of something. An explicit description is important since sensemaking is a complex cognitive process in itself[21;41].

I address this gap in the literature by probing the association between modeling and sensemaking by qualitatively analyzing the case studies of two students, Matthew and Ken, sensemaking about a physics problem by modeling the given context. The participants' modeling is examined by noting their construction of mental models[42], and their subsequent expression of the model's ideas using Suárez's Denotative Function (DF), Demonstration (D), and Inferential Function (IF) – the DFDIF account of modeling[4]. Additionally, the participants' sensemaking is examined through Odden and Russ's Sensemaking Epistemic Game[40]. The case studies reveal that particular aspects of modeling and sensemaking processes cooccur. For instance, the priming on the 'given' information from the problem statement constituted the students' engagement with their mental models, and their attempts to resolve inconsistencies in understanding involved the use of external representations. I also note that when the barriers to modeling are experienced, sensemaking is inhibited.

This work makes six key contributions to the contemporary literature:

1. Leveraging contemporary and cross-disciplinary theoretical ideas in developing physics assessments.
2. Adopting an agent-based approach in physics task-design.
3. Introducing Conjecture Mapping to the context of assessment design in physics.
4. Introducing a methodology extendable to unpack task features which can elicit other valued epistemic practices (such as modeling and argumentation).
5. Unpacking the intertwining between modeling and sensemaking processes.
6. Identifying modeling barriers as one of the factors which can inhibit students' sensemaking.

This dissertation is structured as follows: in the next chapter (Chapter 2), I present a brief overview of the literature on modeling and sensemaking processes. In Chapter 3, I discuss the theoretical development of TaSPA before exploring the features of assessment tasks that promote sensemaking in Chapter 4. In Chapters 5 and 6, I analyze students' response to tasks embodying the proposed features. In the end (Chapter 7), I conclude by summarizing the work and discussing the avenues for future research.

Chapter 2

Literature review

In this chapter, I present a review of the literature on modeling and sensemaking processes. A broad overview of the literature indicates that researchers have investigated these two constructs by defining them in a variety of ways. In the following subsections, I will attempt to paint a picture of the broad trends in the literature as it relates to our current study. For a fuller account of the state of research into sensemaking and modeling, references[21;43] and references[44;45] are recommended.

I find four main categories across which modeling has been studies in science education research: (i) analysis of students' mental models, (ii) modeling during problem solving, (iii) modeling in classroom instruction, and (iv) modeling in laboratories. On the sensemaking front, I note three broad categories: (i) theoretical accounts, (ii) analytical accounts, and (iii) cognitive outcomes of sensemaking. The categories across modeling and sensemaking have been detailed below.

2.1 Modeling

Model-based reasoning has gained considerable attention in PER especially in the last three decades. The increasing emphasis on modeling is evident from 'Developing and Using Models' being considered as one of the key scientific practices to be promoted in classrooms[13;46].

The literature on modeling in PER can be mainly classified into: analyzing students' mental models[42;47–51], modeling during problem solving[52–56], promoting modeling in classroom instruction[20;52;53;57–65], and modeling in laboratories[66;66;67;67–72].

Investigations into mental modeling have primarily focused on students' construction and deployment of mental models. Greca and Moreira[42] define mental models to be *'an internal representation which acts as a structural analogue of situations or processes'*. Researchers have observed students to possess several simultaneous mental models whose probability of deployment depends on various factors including student's mental state, assessment statements, instructional methodologies, and peer interactions[48;50].

Problem solving has been one of the primary contexts in which researchers have discussed modeling in physics education[52–56]. These studies have considered modeling in general, and mathematical modeling in particular to be an integral part of problem solving in physics. Researchers have noted several factors such as description of the target system, and the interactions between the system's components to guide the choice and construction of appropriate models during problem solving[54;56].

The third domain of investigations on modeling – promoting modeling in classrooms (part of which overlaps with the studies on modeling during problem solving) – has been advocated to emphasize the application of physics theories and principles to real-world contexts[52;53]. To promote this practice through teaching, models have been classified as mathematical models[52;56], physical models[42], conceptual models[53], and models of objects, systems, interactions and processes[58]. Brew and colleagues have extended the Modeling Instruction to university level by explicitly focusing on representations during classroom discourses[20;60;61;63–65]. This explicit emphasis on modeling has been found to positively impact self-efficacy of women in learning physics[60], facilitate equitable learning environments for students[61], cause positive attitudinal shifts towards physics learning[63;64] and promote students' use of representations during problem solving[65].

In addition to promoting modeling in classrooms, investigations in PER have also emphasized modeling during laboratory experiments. Reasoning around models in laboratories has been advocated since models act as the connecting link between theory and experimentation[67]. The call for designing model-centered laboratory activities has been addressed through development of frameworks and assessments[66;68;70]. Zwickl *et al.*[72] have noted that an explicit emphasis on models in laboratories can enhance the blending of conceptual and quantitative reasoning during experimentation.

In the landscape of modeling literature in PER, I aim to place this study in the categories of students' construction of mental models and modeling during problem solving. In the sections to follow, I discuss two case studies of students solving a physics problem by initially constructing a mental model and subsequently reasoning by constructing representations in order to make sense of the given task.

2.2 Sensemaking

Science education literature has a rich repository of investigations on students sensemaking about their surrounding world. These explorations broadly span across three domains: (i) theoretical descriptions of the sensemaking process (ii) analytical accounts exploring approaches of sensemaking, and (iii) the outcomes of sensemaking. I present a brief overview of the studies in each domain, and encourage readers to go through references[21;43] along with the cited literature for additional details.

2.2.1 Theoretical accounts of sensemaking

A section of the sensemaking literature has focused on theorizing the underlying process involved in 'making sense' of a given context. These accounts have explored sensemaking through the lens of transfer[41], modeling[27;35;38;39;46;73–75], argumentation[23;76;77], epistemic frames[78;79], and epistemic games[40].

According to Nokes-Malach and Mestre[41], sensemaking is a critical component of 'transfer' – the process of leveraging existing knowledge in addressing novel problems. The authors argue sensemaking (during problem-solving) to be an iterative process involving coordination between prior knowledge and contextual information while generating an optimal solution.

The process of narrowing down on the optimal solution is often achieved by 'modeling' the given problem[27;35;38;39;46;73]. Modeling as sensemaking entails an initial construction of mental models, and subsequent validation of the models' ideas through external representations[74]. One can also model the given context by employing mathematics as either a tool and/or as an object of investigation[75]. Choosing the optimal solution candidate is also achieved through construction and critique of claims during an argument[23;76;77]. Sensemaking from the argumentation perspective entails generation and evaluation of new knowledge, both at the individual, and at the community level.

The idea of generating new knowledge is also resonated in other studies theorizing sensemaking as an 'epistemic frame' (a tacit understanding of 'what's going on here?'[17;80]), or as an 'epistemic game' (a strategic approach in perceiving an inquiry[81;82]). Sensemaking epistemic frame involves generating new explanations in response to a perceived gap in one's understanding about an observed phenomenon. These explanations are based on one's lived experiences, and often are aimed at unpacking the underlying mechanism that gives rise to the phenomenon[25;79;83;84]. On the other hand, Sensemaking Epistemic Game[40] conceptualizes sensemaking as a multi-stage iterative process with a goal of addressing of one's knowledge gap by leveraging available information and existing ideas.

2.2.2 Analytical accounts of sensemaking

A faction of the science education literature has focused on analytically identifying reasoning approaches, or instances (mainly involving mathematics) that qualify as sensemaking[25;29;85–92]. A review of this literature reveals a variety of definitions adopted to analyze the sensemaking process.

A subset of this literature has defined sensemaking as establishing coherence between multiple representations of physics knowledge[25;85;86]. Multiple representations such as equations, figures, tables, or linguistic phrases are frequently employed in conveying contextual information in physics. While Emigh *et.al.*[85] define coordination between these forms of representations as sensemaking, Lenz *et.al.*[86] observe sensemaking as seeking coherence or meaning between them.

Other studies have defined sensemaking as an effort to establish connections between the structure of mathematical formalisms, and the physical world[29;87–92]. These studies have observed ‘mathematical sensemaking’ to entail mapping of causal relations[87], conceptual understanding[29;88], or intuitive reasoning[89] about physical systems with formal mathematics.

2.2.3 Cognitive outcomes of sensemaking

The third category of the sensemaking literature focuses on probing the cognitive outcomes of the sensemaking process. This literature posits three major outcomes of sensemaking: generation of new knowledge, development of sophisticated epistemology, and enhanced content understanding.

Generation of new knowledge

As noted in Section 2.2.1, one of the characteristics of sensemaking is the generation of new knowledge by blending curricular ideas with lived experiences. Studies discussing episodes of sensemaking have noted students making novel claims by constructing analogies, making assumptions, designing thought experiments, and predicting outcomes[93–96]. Furthermore, sensemaking also entails a crucial component of scientists’ and engineers’ reasoning in knowledge construction while solving cross-disciplinary real-world problems[25;41;94].

Sophisticated Epistemology

Personal epistemology - perspectives about what it means ‘to know’, and the nature of knowledge plays a crucial role in how one engages with a given task[97–99]. While sensemaking, students iteratively coordinate and reconcile between different forms of knowledge and reasoning approaches. The knowledge forms include lived experiences, intuitive arguments, conceptual and procedural ideas, or hypotheses[21;25;93]. These knowledge forms are further accompanied with reasoning practices such as argumentation[23], asking questions[26], or engaging with models[74;100]. This virtue of leveraging different forms of knowledge and practices, or shifting between the different 'epistemic frames' during sensemaking results in sophisticated epistemology[16;18].

Enhanced content understanding

One of the consequences of generating new knowledge through sophisticated epistemology is enhanced content understanding. Leveraging distinct sources of knowledge and epistemic practices contributes towards better content understanding[17;19;101–104] by equipping students to 'transfer' skills across multiple disciplines[22;41;94].

Chapter 3

Theoretical Approach for Promoting Scientific Practices Through a Scalable Standardized Assessment for Undergraduate Physics

In this chapter, I elucidate the design process of an RBA under development - the Thermal and Statistical Physics Assessment (TaSPA). The goals of the TaSPA include collecting data on students' application of existing knowledge to novel contexts ("knowledge-in-use"), generating actionable feedback for the course instructors, and having the instructor use that feedback to modify their course. Below, I describe the frameworks that motivate the design of the TaSPA, but offer a brief summary of the process here to guide the reader. When available, instructors will be able to use the TaSPA portal to assess their students' learning in thermal and statistical physics by selecting a set of valued (knowledge-in-use style) learning goals from a list. Through a web-link accessed by an email from the online portal, students can take the TaSPA online, making the assessment available even beyond the classroom hours. After a stipulated time, instructors would receive a report on their students' performance on the assessment along with actionable feedback (if required) about how they can modify their courses to better support their students' learning.

In the rest of this chapter, I describe the theories underlying the development of the TaSPA and then discuss how I use them to develop tasks and provide actionable feedback for instructors. In doing so, I address the research question:

How can a scalable, research-based assessment for upper division thermal and statistical physics be developed that assesses "knowledge-in-use", and provides actionable feedback for instructors to support their students' learning?

In the next section, I present the "theory-of-action" before discussing the frameworks and their operationalization in the assessment development.

3.1 Theory-of-action for TaSPA

The process of assessment design typically involves articulating the measurement argument, i.e., what the assessment intends to measure (e.g., conceptual understanding), and the mechanism through which this measurement occurs (the ways of scoring and interpretation of the scores to deduce inferences about conceptual understanding)[105]. The measurement arguments assume significance as they implicitly communicate what the education research community considers important to assess, and can influence instructors on guiding the everyday activities in classrooms.

For the current work, in line with the contemporary discussions in the assessment literature, I consider assessments also as "instruments of change"' rather than solely as "instruments of measurement". This paradigm shift then requires replacement of the measurement argument with a "theory-of-action" - an explicit argument on how the assessment brings about the intended change in an individual or an institution[106]. A theory of action entails articulating the components of the assessment, the change the components intend to produce (intended effects), and the mechanism which facilitates this change.

Our intended effects correspond to shifting the emphasis of our classrooms from pure conceptual understanding to application of students' existing knowledge in making sense of novel contexts. The assessment components involve the learning performances (goals), knowledge-

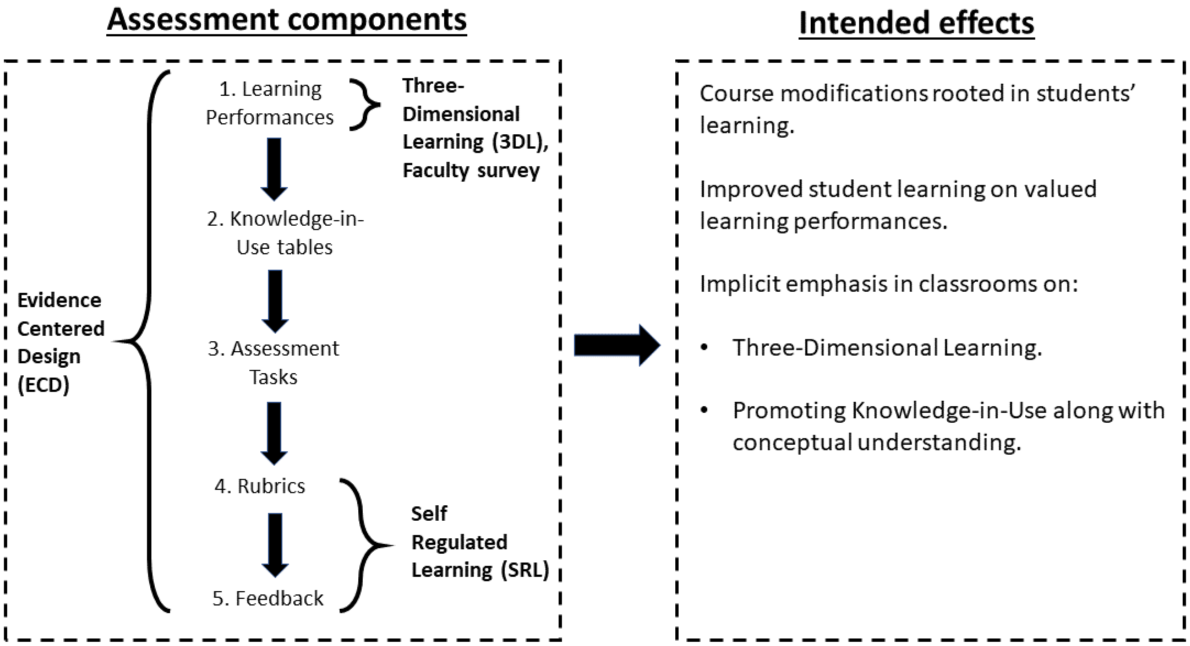


Figure 3.1: *The theory-of-action for the Thermal and Statistical Physics Assessment (TaSPA) highlighting the assessment components (on the left) and the intended effects (on the right). The arrow in between signifies faculty's uptake of our generated feedback.*

in-use tables, assessment tasks, rubrics, and feedback reports. Instructors taking up the generated feedback and making relevant course modifications correspond to the mechanism through which the assessment components lead to the intended effects. In the following section, I describe the three frameworks: Three-Dimensional Learning (3DL), Evidence-Centered Design (ECD), and Self-Regulated Learning (SRL) which are leveraged in the design of TaSPA and, together, form the basis for our theory-of-action.

Figure 3.1 summarizes TaSPA's theory-of-action. The target areas of the assessment are identified from the 3DL and from the physics education literature[11]. The assessment tasks are then developed along with the associated rubrics and feedback reports by leveraging ECD and SRL as shown in the figure. The operationalization of these frameworks in the development of the assessment components is to promote knowledge-in-use through emphasis on scientific practices (along with core ideas and cross-cutting concepts) in classrooms.

3.2 Design and Theoretical Frameworks

The rest of this section entails the description of three frameworks - Three-Dimensional Learning, Evidence Centered Design, and Self-Regulated Learning - which are leveraged in the development of TaSPA's assessment components.

3.2.1 Three-Dimensional Learning

Educators have emphasized the need to shift the focus of academic learning environments from "knowing science" to "doing science", i.e., moving away from conceptual understanding to organizing ideas and engaging in practices in similar ways as disciplinary experts during knowledge construction[107]. Three-Dimensional Learning (3DL) is the design framework underlying the Next Generation Science Standards[13] which promotes the spirit of "doing science" at K-12 level. This framework characterizes students' learning across three "dimensions" namely (i) Scientific Practices - the ways or activities through which scientists generate new knowledge. Examples of these include constructing explanations and engaging in arguments based on evidence, developing and using models, etc. (ii) Core ideas- the key concepts of a discipline which provide organizational structure for students' existing knowledge in addition to facilitating sensemaking of new knowledge. These include energy, properties of matter, etc. (iii) Cross-cutting concepts- the concepts which have applications across multiple scientific disciplines. These include cause and effect, patterns, etc.

3DL places emphasis on blending the "three dimensions" into science curriculum, instruction, and assessment in order to effectively transform classrooms into authentic science learning environments. While this framework was initially proposed for the K-12 education system, researchers have argued for its relevance to the college level[108]. This framework is adopted in crafting the learning performances by blending the thermal and statistical physics content with scientific practices, core ideas, and cross-cutting concepts (Refer Figure 3.1 and 3.2.4). While scientific practices and cross-cutting concepts are directly derived from the existing framework, I add the core idea of "entropy" to better suit the upper division thermal and statistical context.

3.2.2 Evidence-Centered Design

Assessments convey information about students' proficiency on a topic (conceptual understanding) or skill (problem solving), based on their responses to specific tasks. Evidence-Centered Design[30] is a framework which conceptualizes assessment as an evidentiary argument, i.e., an argument which infers about what students can do based on the evidence demonstrated by students in a given learning environment. This framework articulates the assessment design and implementation in terms of the five "layers", namely:

1. *Domain analysis.* The gathering of substantive information about how disciplinary knowledge is contextually acquired and applied.
2. *Domain modeling.* Organizing and relating information obtained in the domain analysis into a coherent organizational structure. This structure specifies what the assessment intends to measure, the evidence required to make claims about students' learning, and details about the assessment task features.
3. *Conceptual assessment framework.* The specifications about the 'nuts and bolts' of the assessment machinery such as details of the task formats, rubrics for evaluation, etc.
4. *Assessment implementation.* Preparation of the elements specified in the conceptual assessment framework.
5. *Assessment delivery.* The operationalization of the assessment such as students taking the tasks, analysis of their responses and generation of feedback reports.

The ECD framework is adopted in developing our assessment components as highlighted in Figure 3.1 by mainly drawing on the work from Harris *et al.*[12]

3.2.3 Self-regulated Learning

"Learning" is a common objective around which researchers, instructors, and students engage with assessments. The notion of learning as it relates to assessment scores has traditionally been focused on students. In this work, instructors are also positioned as learners, or, more specifically, as self-regulated learners who seek to modify their pedagogical practices to maximize their students' learning.

Self-regulated learning[31] is a meta-cognitive activity in which learners regulate aspects of their thought processes, motivation, and behavior during learning. This can include: actively monitoring progress towards goals; consistently analyzing strategies employed in pursuit of the set goals; managing emotions, efforts, resources, etc.; and responding to external feedback. Effective feedback, particularly external, has been considered as a catalyst in accelerating and providing necessary impetus for self-regulated learners in making progress towards their goals. Sadler[109] put forward three features of external feedback that are effective at assisting self-regulated learners in their progress towards the set goals. I adopt these features in providing external feedback for instructors and the features include clarifying: (i) the desired performance, (ii) the current state of the performance, and (iii) opportunities to close the gap between current and desired performance.

The next section details the operationalization of the above-mentioned frameworks in the development of assessment components before discussing the hypothesized mechanisms in producing the intended effects.

3.2.4 Development of TaSPA's assessment components

The assessment components of TaSPA include: (i) Learning performances, (ii) Knowledge-in-use tables, (iii) Assessment tasks, (iv) Rubrics, and (v) Feedback reports. The interplay of the three frameworks discussed in the previous section in the development of these components are highlighted in Figure 3.1 and described below.

1. Articulating learning performances.

Learning performances (LPs) are assessable, knowledge-in-use statements which reflect what students should know, and be able to do with that knowledge[12]. In this way, learning performances reflect the acquisition and application of disciplinary knowledge in the context of assessment development, and thus mark the "domain analysis" layer of the Evidence

Centered Design. The focal content areas of the learning performances are drawn from Rainey *et al.*'s survey[11] documenting the conceptual ideas valued by instructors who have recently taught thermal/statistical physics courses across the United States. I particularly focus on the ideas identified by 85% of the respondents (or higher) as valuable to be assessed in their classes.

The process of crafting a learning performance includes an initial articulation of a sub-idea - a conceptual statement encompassing one or more content areas from the survey. Below is an example of a sub-idea which addresses the survey's topics of 'heat', 'temperature', and 'work':

> "Heat flow changes (or does not change) temperature depending on the process undergone and ambient conditions (e.g., pressure, temperature)".

The sub-idea is then blended with the three elements of 3DL, i.e., scientific practices, core ideas, and cross-cutting concepts leading to a learning performance. Though the scientific practices and the cross-cutting concepts are adopted directly from the K-12 framework[46], the core ideas were modified to better suit the upper-division context. The core idea of "Entropy" is added to the existing pair of "Energy", and "Matter and its Interactions" to better capture the conceptual terrain of undergraduate thermal and statistical physics. This choice (of adding a core idea) aligns with suggestions made for extending the 3DL framework for the college level physics courses[108]. Below, I present one of the learning performances associated with the above-mentioned sub-idea. The associated scientific practice ("Constructing Explanations") has been highlighted by the bold text, the core idea ("Energy") with underlined text, and the corresponding cross-cutting concept ("Cause and effect: mechanism and explanation") with the italicized text.

> By the end of my course, students should be able to: **Generate an explanation** about the *mechanism by which the temperature does (or does not) change* with <u>heat flow into or out of a system</u> informed by the process undergone and ambient conditions (e.g., pressure, temperature).

The choice of the exemplar learning performance and the associated assessment components (discussed below) is guided by its accompanying simplicity in demonstrating the development process within the available space constraints.

2. Generating knowledge-in-use tables.

The next step involves articulating the knowledge, skills, and abilities required to demonstrate proficiency in addressing the learning performance. This is followed by the articulation of the corresponding evidence statements (ESs) - the evidence in students' work which reflect the required proficiency. Lastly, the task features which can engage students in the targeted scientific practice, core idea, and cross-cutting concept are identified. The task features are drawn from the Three-Dimensional Learning Assessment Protocol (3D-LAP) which lays out the criteria for tasks to elicit evidence of students' abilities to engage with the "three dimensions" of learning[108].

The organization of the above-mentioned elements, i.e., the learning performance, the associated knowledge, skills and abilities required to demonstrate the proficiency in the learning performance, the corresponding evidence statements and the task features in a tabular form is referred to as the "knowledge-in-use" table. Our approach in developing this tabular structure is strongly influenced and guided by the work from Harris *et al.*[12]. Table 3.1 presents the knowledge-in-use table for our exemplar LP. From the ECD framework's perspective, the formulation of this table corresponds to the "Domain modeling" layer.

3. Developing assessment tasks

The task development begins by identifying a context which addresses each of the elements in the Knowledge-in-Use table (Table 3.1). The identified context is then employed in developing an open-ended task. This task is then piloted to upper-division undergraduate physics students, and their responses are analyzed through the lens of the evidence statements. The student responses assist in making necessary modifications to the knowledge-in-use table in addition to assisting the development of a coupled multiple-response (CMR) version of the

Learning performance	Generate an explanation about the mechanism by which the temperature does (or does not) change with heat flow into or out of a system informed by the process undergone and ambient conditions.
Knowledge, Skills and Abilities	(i) Identifying relevant processes and conditions that influence temperature changes, and (ii) Describing the physical mechanism justifying why heat flow does or does not change the temperature.
Evidence Statements	(i) Statements about processes and conditions that influence temperature changes and (ii) Statements describing the physical mechanism justifying why heat flow does or does not change the temperature.
Task Features	(i) Providing a claim about a phenomenon, and (ii) Asking students to reason about the claim by accounting the cause and effect using appropriate scientific principles.

Table 3.1: *An exemplar knowledge-in-use table.*

task[110]. A coupled multiple-response task is a form of two-tier multiple-choice question that consists of two forms of questions: (i) multiple choice (having a single correct response among the provided options) and multiple response (having multiple correct or partially correct responses). Though not as effective as open-ended prompts in capturing detailed insights on students' reasoning, the CMR formats provide advantages in streamlined/automated scoring, thus facilitating the "scalability" feature of the TaSPA. For additional details on creating CMR versions from the responses to the open-ended prompts, refer[111].

4. Designing rubrics

Rubrics translate students' responses to the assessment tasks into scores or ratings which in turn are interpreted with respect to ideal outcomes. Since the tasks and rubrics correspond to the "nuts and bolts" through which the "assessment machinery" functions, their development marks the "Conceptual assessment framework" layer from the Evidence Centered Design's perspective (Refer 3.2.2). The rubrics are designed to reflect the extent of evidence provided in students' responses demonstrating the extent to which I can claim they have achieved the given learning performance. Students' responses are characterized into strong, weak, or no evidence which I refer to as their proficiency "Met", "Partially Met" or "Not met" reflecting their knowledge, skills, and abilities in meeting the learning performance. The task development (particularly the CMR version) and the rubric design often occur simultaneously to ensure coherence in capturing relevant evidence for students' proficiencies. Every response pattern to CMR tasks is associated with a rubric rating which characterizes the aforementioned proficiency levels.

5. Generating Feedback Reports

One of the key objectives for the TaSPA is to provide actionable feedback to instructors. I position instructors as self-regulated learners who constantly modify their teaching strategies based on personal or systemic feedback. In order to ensure uptake, our feedback embodies the effective principles as specified in Section 3.2.3. Additionally, accounting the instructors' diverse classroom practices across institutions, the feedback refrains from providing specific, prescriptive approaches. Instead, I focus on the nature of opportunities that students can be provided with, to better support their students' learning. However, the ways of operationalizing these opportunities are left to the instructors.

The process of generating feedback involves translating and interpreting individual students' responses to make inferences about the class' performance. Though there are several ways to achieve this objective, Voting Theory[112] is currently being explored to guide our choices. The generated feedback reports based on the class' performance are rooted in the rubrics and are categorized based on the evidence statements. Every rating of the rubric accompanies a feedback statement which specifies the desired performance, students' current performance, and the opportunities which can bridge the gap between the desired and the current performance (if required).

In our example, if students identified the relevant processes or conditions (evidence statement 1) without completely accounting for the underlying mechanism for the changes in temperature (evidence statement 2), it would reflect "proficiency met" criteria for evidence statement 1 and "proficiency partially met" for Evidence Statement 2. In such a case our feedback would look like:

You wanted to assess if your students are able to: Generate an explanation about the mechanism by which the temperature does (or does not) change with heat flow into or out of a system informed by the process undergone and ambient conditions.

Students were asked to:

- Identify relevant processes and conditions that influence temperature changes.
- Describe the physical mechanism justifying why heat flow does or does not change the temperature.

The TaSPA:

- Provided evidence that your students identified relevant processes and conditions that influence temperature changes.
- Provided evidence that your students partially accounted for the mechanism justifying why heat does or does not change the temperature of the system.

Students could benefit from more opportunities:

- To explore the physical mechanisms through which the heat flow into or out of the system affects its temperature.

This feedback structure was designed with attention to both simplicity and actionability. To maximize the chance the instructor will process and act on the feedback, I wanted to ensure the presentation was not overwhelming. However, as the feedback will be delivered online, options will be built in for instructors to get more fine-grained suggestions if they are interested.

3.3 Intended effects of the TaSPA's feedback

Assuming that instructors find the generated feedback clear and actionable, it is reasonable to expect that they would then initiate changes in classroom instruction that could result in improved student learning. Going back to TaSPA's theory-of-action (Refer Figure 3.1 and 3.1), our first intended change would be on students' learning on the valued LPs. Since the generated feedback is rooted in evidence statements, which are in turn derived from the learning performances, any uptake of feedback is also expected to result in students' improved learning on the corresponding learning performance. Furthermore, since the learning performances are crafted by blending scientific practices, core ideas and cross-cutting concepts, I posit that corresponding course modifications based on our feedback might also result in implicitly promoting three-dimensional learning and thus knowledge-in-use in science learning environments.

3.4 Conclusion

I present a broad overview of the development process of a novel research-based assessment - the Thermal and Statistical Physics Assessment. Specifically, I elucidate the leveraging of existing theoretical and design frameworks in the development of TaSPA. This work presents a paradigm shift in how assessments are envisioned and designed by the discipline-based education research community. I have presented our "theory-of-action"- an explicit account of how our assessment can bring about an intended change by shifting the focus of our classrooms from "knowing" to "doing" science.

Furthermore, the contemporary literature has raised several concerns on the existing assessments which include (i) lack of clarity for instructors in interpreting students' assessment scores, (ii) the need to shift the focus of assessments from what students know to what students can do with what they know (knowledge-in-use), (iii) misalignment between the assessment goals and the valued objectives of the instructors, and lastly (iv) scarcity of standardized assessments for undergraduate thermal and statistical physics. TaSPA addresses the first concern by providing actionable feedback for instructors on how they can go about modifying their courses based on their students' responses to the assessment tasks. Additionally, by blending scientific practices with the conceptual ideas of thermal and statistical physics, TaSPA focuses on assessing the 'doing' aspect of science, thereby addressing the second concern. Since instructors get to choose what they wish to assess (by choosing the learning performances they value), the current assessment contributes in minimizing the existing gap between the objectives of assessment and local goals of instructors (thereby addressing the remaining two concerns).

Currently, the assessment stands with a total of thirteen learning performances along with the associated components (tasks, rubrics and the feedback reports). In the process of developing and finalizing these components, a total of 176 student responses across two large R1 universities in the United States were analyzed. Furthermore, in order to get instructors' perspectives on the assessment, particularly the learning performances and the feedback statements, ten faculty members who were teaching or had recently taught thermal and statistical physics at various institutions across the United States were interviewed. The faculty interviews mainly revealed an overall positive response on the learning performances reflecting the viability of scientific practices, core ideas, and cross-cutting concepts for undergraduate level. Secondly, every faculty member considered a different learning performance as important, thereby reinforcing the need for a design feature allowing instructors to select assessment objectives which they seem valuable to be assessed.

As for the future work, theoretical underpinnings in translating and interpreting the individual responses to the class' results are being explored. The process of piloting the entire exercise of our assessment administration is also underway, starting from instructors choosing learning performances to getting actionable feedback based on their students' responses at numerous undergraduate institutions in the United States. The assessment is being planned to be available to the public in 2024.

Chapter 4

Theoretical exploration of task features that facilitate student sensemaking in physics

In this chapter, I present theoretical arguments in favor of four task features that facilitate student sensemaking in physics.

Researchers in physics education have advocated for facilitating students' content understanding through promoting "sophisticated epistemology" - leveraging different modes of reasoning while engaging with a task[15;16]. The education research community has also emphasized promoting pedagogical practices that facilitate students in generating new knowledge by building on their existing ideas[20]. Sensemaking[21] – the process of addressing a perceived gap in one's understanding – attends to these valued objectives.

Sensemaking assists students' in better comprehending the curricular content by leveraging different forms of knowledge and practices[22]. Sensemaking is also one of the many ways through which scientists and engineers generate new knowledge[23–25]. Given this significance, there has been an uptick in investigations on the discourse markers and the nature of tasks associated with sensemaking. These include (but are not limited to) construction and critique of claims[23], vexing questions during interactions[26], the blending of model- and evidence-based reasoning[27], computational reasoning about physics scenarios[28], addressing quantitative problems through qualitative insights and vice-versa[29], and explaining physical systems through mathematical insights[113]. I contribute to these efforts by theoretically exploring the assessment task features that increase the likelihood of students sensemaking in physics.

I identify these features by initially noting characteristics of the sensemaking process as described in the science education literature[21]. Guided by the research in cognitive psychology, science education and philosophy of science, I make a theoretical argument for the task features that promote sensemaking in physics. I neither argue that the proposed features *necessarily* engage students in sensemaking nor any task that elicits sensemaking *necessarily* entails these features. I also do not claim that the proposed list is an *exhaustive* one. Rather, I make a modest argument that tasks entailing the proposed features *together* (as opposed to presence of one of these) increase *the likelihood* of students sensemaking.

To highlight how the proposed features bring about the desired outcome of sensemaking, I elucidate the design criteria through a conjecture map. Conjecture mapping is a framework primarily employed in design-based research to conceptualize the interactions between theoretically salient design features of a learning environment and their intended outcomes[1]. I adopt this framework to my context in elucidating how the proposed features elicit the desired outcomes of sensemaking.

The current work makes four key contributions to the contemporary literature. Firstly, this study presents an agent-based approach in articulating the task features by shifting the vocabulary from "*tasks entailing a feature X*" to "*tasks that cue students about X"* or "*tasks that cue students to do X*". Such vocabulary would better account for students' agency along with the local practices of their learning environments. Secondly, this work operationalizes the Conjecture Mapping framework in the context of task-design in physics. Thirdly, I leverage cross-disciplinary theoretical ideas particularly from cognitive psychology and philosophy of science in identifying the task features that promote sensemaking in physics. Lastly, my methodological approach in identifying task features can also be potentially extended in unpacking task features which can elicit other valued epistemic practices such as modeling and argumentation.

In doing so, I address the following research questions in the rest of this chapter:

RQ1 How can I adopt a framework-based approach in theoretically identifying task features that promote students sensemaking in physics?

RQ2 What set of assessment task features increase the likelihood of students sensemaking in physics?

In the next section, I detail the theoretical stances adopted to address the above-mentioned research questions.

4.1 Theory

4.1.1 Agentic paradigm in task-design

Research on task-design has traditionally involved prescribing a set of design features (often backed by analysis of students' responses) which can elicit a targeted response from students. However, studies have increasingly highlighted the role of contextual factors such as local norms of students' community (e.g., teachers, classrooms, and institutions) on what counts as "knowing" or "doing" science[114], students' agency in accessing knowledge sources[115;116], and their in-the-moment framing of the task's expectations[117] as influencing students' engagement with tasks. Efforts on explicit prompting in tasks too have evoked mixed results. While few have noted explicit prompting to enhance students' understanding on domain principles and procedural knowledge[118–120], others have noted them to impede students' intuitive reasoning[121] by selectively emphasizing parts of the presented information[122;123]. As Berland *et al.*[124] note

> "*[...]emphasizing the actions alone can result in rote performance and attainment of skills, rather than student engagement in the rich work of scientific knowledge construction, evaluation, and refinement.*"

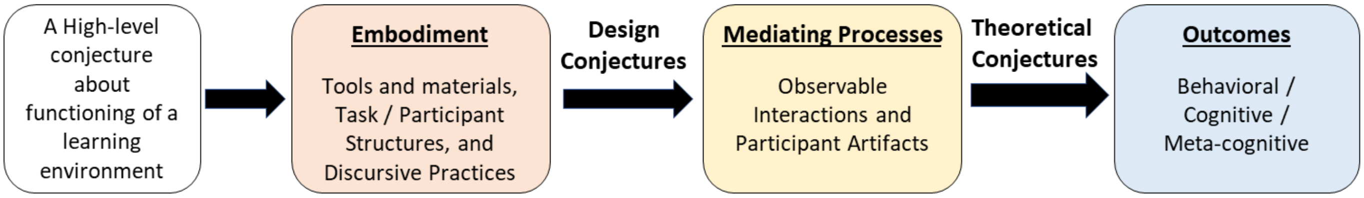


Figure 4.1: *Modified schematic representation highlighting relationships between the elements of a conjecture map. The original representation can be found in*[1].

In light of these observations, I adopt an agent-based approach in arguing about task features by shifting the vocabulary from "*tasks entailing a feature X*" to "*tasks that cue students about X*" or "*tasks that cue students to do X*". By "cuing" I mean, conveying or setting up expectations for students about a feature in a task or about a specific way of reasoning as a solution approach to the task. Such vocabulary would better account for students' agency along with the local practices of their learning environments. In the rest of this paper, I adopt this framing in theorizing assessment features that can increase the likelihood of students sensemaking in physics.

4.1.2 Conjecture Mapping

Design-based research accompanies a set of epistemic commitments about design and functioning of learning environments in addition to advancing the understanding of teaching and learning processes[1]. Attending to these commitments often requires researchers to articulate conjectures about how the designed learning environment functions in an intended setting. Conjecture mapping[1] is a technique which conceptualizes these arguments by establishing relationships between the design features, processes enacted by participants engaging with these features, and the intended outcomes. This technique highlights the relationships between various aspects of educational design through six elements: (i) a high-level conjecture, (ii) embodiment, (iii) mediating processes, (iv) outcomes, (v) design conjectures, and (v) theoretical conjectures (Figure 4.1).

A *high level conjecture* forms the first element of a conjecture map which articulates the theoretical idea driving the design of a novel learning environment. The articulated conjecture provides the road-map of the theoretical idea's operationalization in a given setting. This conjecture is then reified in *embodiment*, the second element of a conjecture map, which crystallizes the design features into several components. These components include: tools and materials (assessments, devices, etc.), task structures (the nature and form of tasks), participant structures (roles and responsibilities of participants), and discursive practices (forms of participants' discourses). These components further contribute to the *mediating processes*, a set of interactions and artifacts produced from the participants that mediate between the designed features and the intended cognitive/meta-cognitive *outcomes*.

The *embodiment, mediating processes,* and *outcomes* are connected through *design* and *theoretical conjectures* - the last two elements of a conjecture map. Design conjectures are the arguments about how the components of embodiment (tools/materials, task/participant structures and discursive practices) lead to the mediating processes. Theoretical conjectures, on the other hand, are the arguments describing how the mediating processes will in turn result into the desired outcomes. Figure 4.1 schematically represents the elements of a conjecture map and their interrelationships.

I adopt conjecture mapping to elucidate how a set of task features (embodiment) can nudge students to engage "sensemaking elements" (mediating process) leading to generation of new knowledge, sophisticated epistemologies, and enhanced content understanding (outcomes). The theoretical arguments in favour of these outcomes (theoretical conjectures) are discussed in Section 2.2.3. Sections 4.3 to 4.6 detail the arguments about how the proposed task features elicit the features of sensemaking (design conjectures). Figure 4.2 represents the adoption of conjecture mapping to my study.

4.1.3 Sensemaking

Studies in science education have described sensemaking in diverse ways. In the rest of this paper, I adopt Odden and Russ'[21] synthesized account of sensemaking as:

> *a dynamic process of building or revising an explanation in order to 'figure something out' - to ascertain the mechanism underlying a phenomenon in order to*

> *resolve a gap or inconsistency in one's understanding. One builds this explanation out of a mix of everyday knowledge and formal knowledge by iteratively proposing and connecting up different ideas on the subject. One also simultaneously checks that those connections and ideas are coherent, both with one another and with other ideas in one's knowledge system.*

Odden and Russ put forward this definition by synthesizing three approaches through which researchers have conceptualized sensemaking in science education. In the first approach – as a stance towards science learning – sensemaking has been noted to entail generation of explanations describing the underlying mechanism of a phenomenon. In the second approach – as a cognitive process – sensemaking has been noted in involve integration of prior knowledge (experiences) with formal knowledge. In the last approach – as a discourse practice – sensemaking has been conceptualized as construction and critique of claims during argumentation. The construction component of argumentation entails proposing and connecting ideas to substantiate a claim. The critique component on the other hand, entails ensuring coherence between various the connected ideas.

Based on the definition of the sensemaking process, and the conceptualizations of sensemaking across the three approaches in the science education literature, I note the following "sensemaking elements" or the set of activities crucial for engaging in sensemaking:

1. Use of everyday and formal knowledge while reasoning about a phenomenon (sensemaking as a discourse practice).
2. Ascertaining the underlying mechanism of the phenomenon (sensemaking as a stance towards science learning).
3. Generating and connecting up different ideas in one's knowledge system (sensemaking as a discourse practice).
4. Seeking coherence between the generated ideas (sensemaking as a discourse practice).

It should be noted that the above elements do not take into account the crucial aspect of noticing inconsistencies in one's understanding during sensemaking. The noticing of a discrepancy in one's knowledge system (also referred to as "gap in one's knowledge system" or "epistemic uncertainity") is a highly contextualized activity influenced by various factors including prior knowledge, awareness, self-evaluation, and adopted strategies while reasoning about a given scenario[125]. My list of sensemaking elements does not include this critical feature due to its highly contextual nature. In order to address this shortcoming in my theoretical approach, I adopt a probabilistic stance ("the task features *increase the likelihood* of students sensemaking") rather than a deterministic one ("the task features *elicit* sensemaking") in my arguments. In addition, it is also worth noting that sensemaking can occur irrespective of "completeness" or "correctness". That is one can engage in the sensemaking process despite failing to conclusively address or incorrectly address the perceived gap.

I blend the above-mentioned sensemaking elements with conjecture mapping framework in identifying task features which promote sensemaking. By definition, if students engage in all of the above-mentioned sensemaking elements during an activity, they are more likely to engage in the sensemaking process. Along the same lines, I posit as my high level conjecture that *a set of task features which elicit the sensemaking elements increase the likelihood of students sensemaking.* My design conjectures correspond to the arguments (articulated in Sections 4.2 to 4.6) which link the proposed task features to the sensemaking elements of sensemaking.

4.2 Task features that facilitate sensemaking

In Section 4.1.3, I identified four "sensemaking elements" or a set of activities which together contribute to the likelihood of sensemaking. These include: blending everyday and formal knowledge while reasoning about a phenomenon, ascertaining the underlying mechanism of the phenomenon, generating and connecting diverse ideas, and seeking coherence between the generated ideas. I posit that the set of task features which elicit these sensemaking elements increases the likelihood of students sensemaking in physics.

In the next four sections, I propose that tasks which cue students about the following to promote sensemaking in physics: (i) the presence of real-world context(s), (ii) to engage

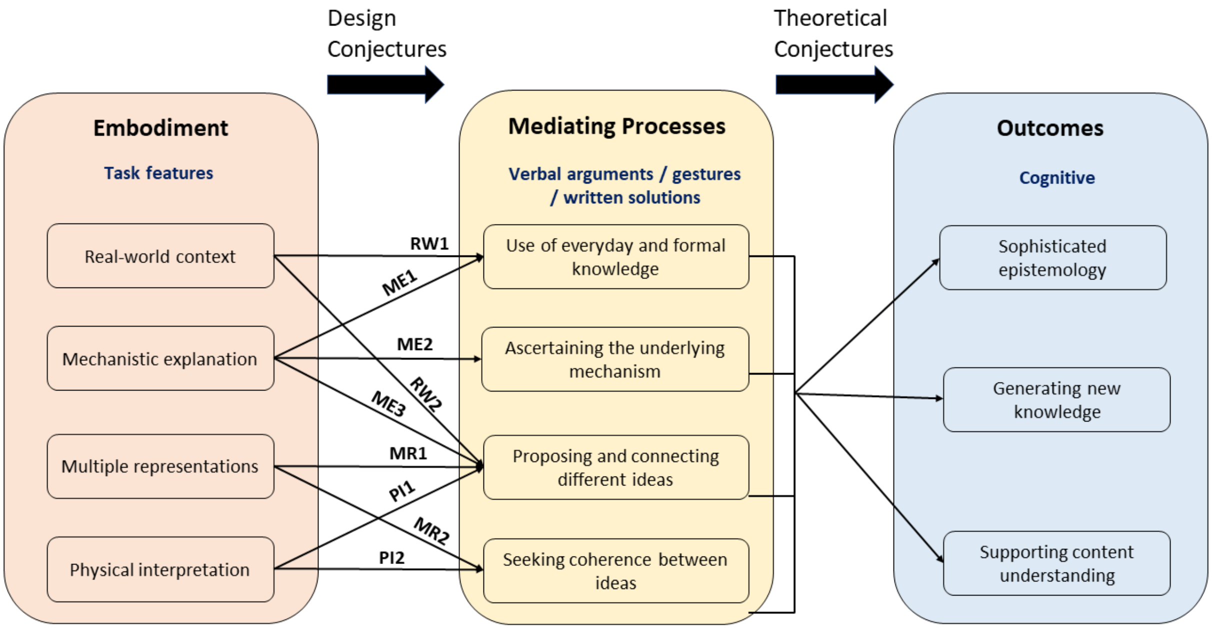


Figure 4.2: *Contextual operationalization of the conjecture map in my study. My high level conjecture (not represented in this figure) takes the form: "a set of task features which elicit the sensemaking elements increase the likelihood of students sensemaking". While the design conjectures are detailed in Sections 4.3- 4.6, the theoretical conjectures are discussed in Section 2.2.3.*

in mechanistic reasoning, (iii) to coordinate between multiple representations, and (iv) to extract physical implications from mathematical expressions. Each section consists of *conjectures* - arguments about a task feature eliciting specific sensemaking elements, *theoretical background* - a theoretical basis of the argument, and *empirical evidence* - evidence in favor of the argument from the literature. Tables 4.1 and 4.2 summarize these components.

4.3 Tasks cuing about the presence of real-world context(s)

The first feature I argue to contribute for students sensemaking is the task cuing students about the presence of real-world context(s). In line with the contemporary discourse in the science education literature, I consider a real-world context as a scenario relevant to the learner, and which requires application of scientific principles/models to make sense of the presented scenario[126]. I argue that tasks perceived as rooted in real-world contexts facilitate two of the four sensemaking elements of sensemaking: use of everyday and formal knowledge; and generating and connecting up diverse ideas in students' knowledge system. These arguments, as design conjectures in my conjecture map (Figure 4.2), have been labeled as RW1 and RW2.

RW1: Real-world contexts facilitate use of everyday and curricular knowledge

Conjecture RW1: *If a task cues students about the presence of a real-world context, then it is more likely to invoke their everyday and curricular knowledge.* In other words, I posit that real-world scenarios in physics tasks appeal to students' lived experiences along with priming their formal curricular ideas. In order to substantiate this argument, I turn to studies in cognitive psychology probing the influence of words or phrases in tasks priming specific information from one's knowledge system.

Theoretical background: Investigations on human interactions with tasks associated with a language's vocabulary (lexical tasks) have observed the role of tasks' contexts on participants' reasoning[127]. According to these studies, the greater the relevance of the task's context to the participants, the better is the task's interaction with their memories[128;129]. 'Semantic priming'[130] is one of the theoretical constructs proposed to explain how words or phrases in a lexical task cue related ideas from one's memory.

Semantic priming is a cognitive effect in which people respond faster to targeted words (e.g., 'dolphin') when they are preceded by related words (e.g., 'whale'), as compared to the unrelated ones (e.g., 'chair'). Semantic relatedness represents the similarity in meaning or the overlap in featural description between a set of words or phrases[131]. Collins and Loftus[132], present the 'Spreading Activation Theory' to describe the mechanism through which semantic memory is accessed during lexical activities. According to this theory, semantic memory consists of a network of interconnected nodes, with each node representing a concept. A "concept" can take several forms ranging from a word to a proposition. The connections between any pair of nodes represent the information connecting the two concepts. The stronger the connections between the two nodes, the easier it is to retrieve associated concepts from memory.

This memory network further embeds a network called semantic networks where the nodes are connected based on the words' meaning, and their shared features. The strength of association between the nodes depends on the degree to which the associated nodes share common features. For instance, the semantic association of the word "red" is stronger with "rose" as compared with "elephant". When a concept is primed during a lexical activity (such as while reading the task prompt), activation spreads out from the primed node along the paths of the network. The "intensity" of activation spread is higher for a strongly associated pair of nodes.

I conjecture that context-based tasks trigger semantic priming with activation spread emanating from concepts (nodes) associated with students' lived experiences as well as with their curricular knowledge. In other words, real-world scenarios in tasks are more likely to invoke arguments from everyday lives and formal knowledge. This cuing is more likely to be semantic in nature, i.e., based on shared features of the words/phrases in the task description.

Empirical evidence: I find empirical evidence for my above conjecture from several studies in PER. Odden and Russ[26], while noting the role of vexing questions in sensemaking, discuss a pair of students' (Jake and Liam) reasoning on a task rooted in real-life. The task inquires about the safety of a car's passengers when exiting the vehicle following a lightning strike during a thunderstorm (with the passengers inside the car being unaffected by the lightning). The students approach the task by blending conceptual arguments about charge distribution with their everyday experiences about the shape of the car door's handle. Similar observations reflecting amalgamation of curricular knowledge with lived experiences can be found in case studies involving context-based tasks from other studies in PER[33;100].

A more direct evidence for my conjecture comes from Enghag *et al.*'s[133] study exploring students' reasoning about a context-rich physics problem. The authors note students initiating their approaches by rephrasing the given prompt based on their lived experiences before referring to the underlying physics principles. The authors highlight references to everyday knowledge as instrumental in students' meaning making, and understanding of the physics involved in the task.

RW2: Real-world contexts facilitate generating and connecting diverse ideas

Conjecture RW2: *If a task cues students about the presence of a real-world context, then it is more likely to lead students to generate and connect diverse sets of ideas (conceptual, procedural and intuitive) from their knowledge system.* I again refer to the literature from cognitive psychology, particularly on search and selective retrieval of ideas from memory[134–136] in support of my argument.

Theoretical background: Nijstad *et al.*[137], propose the Search for Ideas in Associative Memory (SIAM) as a mechanism to describe how ideas get generated while engaging in an activity. According to this account, the internal process of idea generation proceeds through two distinct stages: (i) knowledge activation, and (ii) idea production. In the knowledge activation stage, a search in one's memory networks is triggered by a cue from the contextual features of the task. The structure and function of these memory networks is similar to the networks discussed in Activation Spread Theory in RW1. The memory search initiated by the contextual cue results in the retrieval of an image (idea), whose probability of retrieval depends on the strength of association between the cue and the image. In the second stage, i.e., the idea production stage, the initial image (produced in the previous stage) now acts as the triggering cue, leading to the production of an additional image. This chain of image production – a preceding image acting as a triggering cue for a new image – results into a "train of thought" until the information processing session is terminated. The conditions of termination depend on the nature and outcomes of the activity.

I conjecture that presence of real-world contexts in physics tasks are more likely to trigger a diverse set of ideas from one's knowledge system. From the viewpoint of SIAM, this conjecture can be rephrased as: context-based tasks are more likely to trigger knowledge activation leading to generation of diverse ‘trains of thoughts’. This argument, as a design conjecture in my conjecture map (Figure 4.2) has been labeled as RW2.

Empirical evidence: I find several references in cognitive psychology and science education literature in support of my above argument. While discussing the SIAM account, Nijstad *et al*[137] further note that semantically diverse cues (having diverse featural association between the cues) in a task lead to the generation of diverse set of ideas. George and Wiley[138] note people “rely too heavily on familiar or easily accessible information during idea generation”. Other researchers too have made similar observations on the familiarity of contextual cues stimulating generation of novel ideas[139–141]. In science education, Rennie and Parker[142] document students’ perspectives on solving physics problems based on real-life scenarios. The authors note students referring to context-rich problems as “easier to visualize” as one of the emerging themes in students’ responses.

4.4 Tasks cuing students to generate mechanistic explanation

Mechanistic explanations – descriptions unpacking the underlying mechanism of a phenomenon – have been considered more sophisticated as compared to say, occult or teleological accounts[84;143]. In what follows, I argue that tasks cuing students to generate a mechanistic explanation of a (real-wold) phenomenon to elicit three of the four sensemaking elements. These include: referring everyday and curricular knowledge, ascertaining the underlying mechanism of a phenomenon, and proposing and connecting up different ideas in one’s knowledge system. These conjectures have been respectively labelled as ME1, ME2, and ME3 in Figure 4.2. As the ME2 conjecture – tasks requiring students to generate a mechanistic account lead to mechanistic reasoning – is self explanatory, I will exclude it from detailed discussions below.

ME1: Mechanistic explanations facilitate use of everyday and curricular knowledge

Conjecture ME1: *If a task cues students to generate mechanistic explanation(s), then it is more likely to invoke references to everyday and curricular knowledge.* I substantiate my argument by referring to studies on storage and accessibility of knowledge about mechanisms in memory.

Theoretical background: The cognitive science literature argues for six possible formats through which knowledge about mechanisms (henceforth referred to as 'mechanism knowledge') is internally represented. These include: (i) associations, (ii) forces or powers, (iii) icons, (iv) placeholders, (v) networks, and (vi) schemas. A detailed discussion about each of the representational formats would be beyond the scope of this paper. However, I briefly describe each of these formats, and encourage readers to go through[144–146] along with the cited references for additional details.

"Associations" represent the mapping between two or more distinct events from one's memory such that the knowledge about a familiar event guides the expectations about the unfamiliar one[147;148]. In physics, this association can be observed in the ways propagation of sound in a medium is explained in terms of the compressions and rarefactions occurring on a vibrating spring. The second format, "forces"[149;150] or "powers"[151;152], posits that mechanistic inferences are driven by the knowledge of physical laws. According to the "forces" account, interaction between entities (e.g., collision between two objects) are mediated by forces, and this interaction is described through vectors highlighting the direction of the entities' motion in the influence of the involved forces. On the other hand, the "powers" account posits that humans comprehend mechanisms by conceptualizing entities as having inherent dispositional features. These features either take the form of "powers" (tendency to bring about effects) or "liabilities" (tendency to undergo the effects). Melting of ice in presence of heat, for instance, would be explained in terms of the "power" of heat (causing the ice to melt) and the "liability" of ice (to melt in the influence of heat).

The third candidate – "icons" – is a representation in which mechanisms are conceptualized as mental simulations or mental models consisting of a series of icons or image-like formats[153–155]. The mechanistic imagery that humans possess about the functioning of gears or pulleys is an example of this format[156]. On the contrary, the "placeholders" account (the fourth representational format) posits that people tend to hold a placeholder or a reference pointer for mechanisms instead of a detailed knowledge[157]. Studies arguing for this format have observed people to possess skeletal details about the functioning of familiar everyday complex systems (such as sewing machines or can openers) with a meta-representational placeholder representing an unknown existing mechanism.

The penultimate representational format, "networks", has its origin in statistics and artificial intelligence. According to this account, causal relations are internally comprehended through causal networks (or "Causal Bayesian Networks") in which the nodes represent the variables involved in a mechanism, and the links between the nodes represent the causal relations between the involved variables[158–160]. As an example, the experience of drinking coffee leading to the sense of feeling energized would be represented in a typical causal network with "drinking coffee" and "feeling energized" as two nodes with an arrow pointing from the former towards the latter. The last candidate in my list – "schemas"[161] – correspond to clusters of knowledge in the long-term memory that are employed while figuring out the mechanism of a phenomenon. For instance decisions on the appropriate container to carry cold drinks during summer, are guided by the schemas about heat conductivity through various kinds of materials encountered in daily lives.

One of the common themes across the six representational formats discussed above is their association with one's prior knowledge. The formats highlight that people construct mechanistic accounts by building on their existing notions about the functioning of their surrounding world. Consequently, I posit that tasks cuing students to generate a mechanistic explanation, particularly about a real-world context/phenomenon, can nudge them to invoke their everyday ideas in addition to knowledge gained from formal instruction.

Empirical evidence: Several studies in physics education provide empirical evidence in support of my argument. For instance, diSessa[99] observes students to have a "*sense of mechanism*" through which they gauge the likelihood of various events, make "backward and forward chaining" of events[143], and provide the causal account of an observed phenomenon. This sense of mechanism is built from basic sensemaking elements called "phenomenological primitives" which are in turn derived from one's lived experiences. Resonating a similar view, Hammer[162] notes students and physicists to have "rich stores of causal intuitions", and generating mechanistic explanations to entail references to lived experiences and formal ideas. Sirnoorkar *et al.*[33;100] too observe student-generated mechanistic account of an amusement park ride (a real-world context) to involve an amalgamation of lived experiences and curricular ideas.

ME3: Mechanistic explanations facilitate generation and connection of different ideas

Conjecture ME3: *If a task cues students to generate a mechanistic explanation, then it is more likely to lead them in generation and connection of diverse ideas from their knowledge system.* I support my argument by discussing the nature and features of mechanistic reasoning as described in the philosophy of science and science education literature. To begin with, as noted above, mechanistic reasoning entails drawing ideas from lived experiences and curricular knowledge. Thus, intuitive and formal insights contribute to the spectrum of ideas invoked in unpacking the mechanism of a phenomenon.

Theoretical background: Furthermore, mechanistic reasoning is a complex cognitive process involving description of the behaviour of relevant entities and processes that give rise to a phenomenon[96;143;163;164]. One generates mechanistic accounts by transitioning from observable features of the phenomenon at the macro level to the underlying entities or processes (often at the micro level)[96;143]. The process of ascertaining the mechanism can further involve transitioning back from the micro to the macro features, and testing the validity of the generated explanations by varying the spatial or temporal organization of the entities or processes. This cyclic navigation across "scalar levels" – between observable features and underlying entities or processes, requires one to invoke conceptual, procedural or intuitive ideas and establish coherence between them. This argument as my design conjecture, has been labelled 'ME3' in Figure 4.2.

Empirical evidence Several studies describing episodes of mechanistic reasoning have noted students invoking and connecting diverse sets of ideas in their explanations[143;165;166]. Russ *et al.*[143] discuss first-grade students' mechanistic account of a scenario involving a piece of paper and a book simultaneously dropped from a same height. The students explain the mechanism of falling objects in terms of gravity (a curricular idea) and everyday experience of jumping and landing back on the ground. Similarly, de Andrade *et al.*[165] discuss a pair of middle school students' collaborative exploration of how antacid pills neutralize stomach's acidity. The students (Iris and Raul) generate an explanation by invoking the conceptual argument of the formation of salt and water upon the acid-base reaction. This argument also accompanies a procedural idea of the combination of elements during reaction in determining the molecular formula of the salt and water. The students also reason by making arguments based on everyday experiences that molecules (or objects in general) get smaller in size after collision in a reaction. I find a similar observation in Bachtiar *et al's.*[166] study in which students invoke conceptual, procedural and intuitive ideas while generating mechanistic accounts of a soccer ball's motion while designing its animation.

4.5 Tasks cuing students to engage with multiple representations

Elucidating complex ideas through multiple external representations such as equations (wave functions, equations of state), graphs (kinematic plots, isotherms), or words (laws, theorems) is a common practice in physics. By multiple representations, I mean a combination of distinct external representations that illustrate the same content but use different symbol systems[167;168]. Representational formats of an idea complement each other by highlighting specific information about its content[169–173]. For instance, the kinematic equation $v = v_0 + at$ can better highlight the dependence of an object's final velocity (v) on initial velocity (v_0), duration of its motion (t), and its uniform acceleration (a). On the other hand, the graphical representation of the same equation (velocity vs time plot) better highlights the qualitative variation of the object's velocity for a given nature of acceleration (positive, negative or zero).

I argue that tasks cuing students to engage with multiple representations – either provided or constructed – address the following sensemaking elements: proposing and connecting up different ideas; along with establishing coherence between them. These arguments as my design conjectures, are labelled "MR1" and "MR2" in Figure 4.2. As a primer, note that unlike the last two sections, in the current section and in the next one, I substantiate the the relevant design conjectures through a common theoretical background.

Theoretical background: As a basis for these conjectures, I refer to Mayer's "Cognitive Theory of Multimedia Learning (CTML)"[174] describing the cognitive process involved in interacting with multiple representations. According to CTML, engaging with multiple representations (or multimedia) involves participation of sensory, working, and long term memories. Sensory memory is a short-term memory in which information obtained through sensory inputs (such as visuals of a painting) are stored in their original perceptual form. Working memory corresponds to the cognitive faculty involved in processing and manipulating instantaneous information in active consciousness (e.g., the cognitive process invested in comprehending the meaning of this sentence). Lastly, the long term memory corresponds to the accessible information stored across longer periods of time (e.g., information about one's childhood).

With the participation of these memory forms, the cognitive process involved in interacting with multiple representations proceeds through three distinct and consecutive phases: (i) *selection*, (ii) *organization,* and (iii) *integration* of information. As noted earlier, each representational format of an idea highlights a specific component of the information about the idea. The first phase – *selection* – involves selective choice of this information to be expressed into, or extracted from each representational format with the participation of one's sensory memory. In my kinematic example, the selective extraction of the information about the interdependence of the variables (v, v_0, a, and t), along with their behavior in limiting conditions, mark the *selection* phase associated with the algebraic representation. Similarly, an analogous argument can be made about the graphical representation ($v-t$ plot), in which the qualitative information about the velocity variation is selectively comprehended.

The next phase – *organization* – involves forming mental representations of the embedded, or the selected pieces of information in the working memory. These mental representations are constructed by establishing internal connections between the informational pieces. In the kinematic example, this can correspond to the formation of mental representations of the interdependence of the variables (extracted from the equation), and the velocity variations for a given acceleration (extracted from the graph). Lastly, these mental representations are fused with the help of prior-knowledge drawn from the long-term memory marking the *integration* phase of the CTML. In the kinematic analogy, this phase can correspond to the amalgamation of the algebraic and graphical mental representations using existing knowledge about slopes, or about uniform/non-uniform motion of objects.

MR1: Engaging with multiple representations facilitate generation and connection of ideas

Conjecture MR1: *If a task cues students to engage with multiple representations, then it is more likely to lead students into generation and connection of ideas from their knowledge system.* Based on the CTML's three phases, particularly the *selection* and the *organization* phases, I note that engaging with multiple representations involve generation and connection of ideas. While the former phase involves idea generation through selective interaction with information from the representations, the latter involves connecting the ideas through formation of mental representations.

Empirical evidence: Several studies in physics education have made observations about representations facilitating generation and connection of ideas. Researchers have observed multiple representational formats to cue students in employing and connecting diverse set of domain-specific principles and strategies during problem solving[175–177]. De Cock[175] observes that an isomorphic task presented in varying representational formats tends to elicit different solution approaches along with physics principles. On a similar study, Podolefsky and

Finkelstein[176] note that use of multiple representations can facilitate mapping of ideas during analogical reasoning. Van Heuvelen and Zou[177] note multiple representations of work-energy processes such as verbal descriptions, bar-charts, and mathematical equations facilitate students in better visualizing the energy conservation principle in addition to production of "mental images for different energy quantities".

MR2: Engaging with multiple representations facilitate establishing coherence between ideas

Conjecture MR2: *If a task cues students to engage with multiple representations, then it is more likely to nudge them in seeking coherence between ideas.* Along the same lines, the CTML's last two phases – *organization* and *integration* – highlights that engaging with multiple representations facilitates establishing coherence between the generated ideas. While the former phase entails establishing coherence between the selected pieces of information from a representational format, the latter involves establishing coherence between ideas from representations. Seufert[178] refers to these two phases as 'intra-representational coherence formation' (establishing interrelations *within* a representational format), and 'inter-representational coherence formation' (establishing interrelations *between* representational formats).

Empirical evidence: Cox[172] argues that external representations help in better comprehending an idea as each representational format directs attention to a particular characteristic feature highlighted by the representation. Indeed, Gire and Price[169] note students reasoning in quantum mechanics by effectively coordinating between Dirac, algebraic and matrix notations while representing quantum states of a system. The authors observe students establishing coherence between their ideas by using one notation as a template while creating corresponding representations in other notations.

4.6 Tasks cuing students to extract physical implications from mathematical expressions

Physics education research has an extensive corpus of discussions on the role and use of mathematics in physics[8;179]. A major section of this work has analyzed students' interaction with mathematical formalisms during problem solving[17;19;87]. In the rest of this subsection, I argue that tasks cuing students to extract physical implications from mathematical expressions (equations, plots, etc.) lead to generation and connection of ideas, along with establishing coherence between them. These arguments have been labelled "PI1" and "PI2" in my conjecture map (Figure 4.2).

Theoretical background: Discussions in the philosophy of science literature posit that extracting physical implications from mathematical expressions involve mapping structural features of mathematical formalisms to that of physical systems[3;4;180;181]. This view has been identified with several theoretical perspectives such as 'mapping account'[180], 'interpretation'[3], 'inferential conception'[181] or 'inferential function'[4]. Nevertheless, the underlying theoretical view remains that interpreting mathematical relations involve bridging the structure of mathematical formalisms with the features of the target system. For instance, inferring the motion of a spring (target system) from the equation $F = -kx$ (mathematical structure) involves mapping the algebraic symbol F with the net force on the spring, x with the spring's displacement from its mean position, k with the spring constant and the negative sign with the force's direction. Evidently, this mapping requires one to simultaneously engage with formal mathematical ideas (procedural or conceptual) along with ideas about the physical system. Consequently, I argue that the process of interpreting meaning from mathematical expressions involve generation of ideas and establishing coherence between them.

PI1: Physical interpretations facilitate generation and connection of ideas

Conjecture PI1: *If a task cues students to interpret mathematical expressions in light of physical systems, then it is likely to facilitate generation and connection of ideas.* Physically interpreting mathematical expressions is a common practice in physics. Whether it's determining the likelihood of an event based on the changes in entropy of involved systems, or identifying the position of an image from ray diagrams, students and physicists alike are familiar with this practice.

Empirical evidence: Several studies have noted interpretation of mathematical results as a crucial component of reasoning in physics[55;56;82;104]. Sherin[92] makes a case for the existence of knowledge structures called 'symbolic forms' which mediate the process of meaning making through mathematical formalisms. According to this view, students blend contextual ideas with mathematical insights while interpreting (or expressing) meaning from mathematical expressions. Making a similar observation, Arcavi[182] argues for 'symbol sense' in mathematics, which facilitates interpretation of mathematical expressions via intuitions.

Perhaps, a more direct evidence in support of my argument comes from the study by Kuo *et al.*[90] investigating students' blending of conceptual arguments with formal mathematics. The authors discuss one of their participants' (Pat) reasoning about the difference between final velocities of two balls dropped with differing initial velocities. The reasoning approach involves interpreting a kinematic equation ($v = v_0 + at$) through the lens of derivatives (a mathematical idea), and linking it to the variation of the balls' parameters (ideas of the physical system). Gifford and Finkelstein[75] term this approach as mathematical sensemaking involving use of mathematical 'tools' to reason about physical system.

PI2 Physical interpretations facilitate establishing coherence between ideas

Conjecture PI2: *If a task cues students to interpret mathematical expressions in light of* *physical systems, then it is likely to nudge them in seeking coherence between ideas.* Along the same lines, I argue that the process of interpreting mathematical expressions involves establishing coherence between the generated ideas. In my above example involving the spring's motion, one can interpret the negative sign in the equation as the net force and the displacement vectors being oppositely directed at a given instant of time.

Empirical evidence: Several studies in PER have indeed referred to the process of coherence seeking between mathematics and physical systems as "mathematical sensemaking". While Kuo *et al.* define it as "*leveraging coherence between formal mathematics and conceptual understanding*"[29], Dreyfus *et al.* define the same as "*looking for coherence between the structure of the mathematical formalism and causal or functional relations in the world*"[87]. Wilcox *et al.*[55] further note this practice as 'Reflection of results' while discussing upper-division students' use of mathematics in physics.

4.7 Discussion

4.7.1 Operationalizing conjecture mapping in the context of task-design in physics

I operationalize conjecture mapping – a framework in design-based research – in the context of identifying the assessment task features that promote sensemaking in physics (**RQ1**). Based on the literature's description of the sensemaking process, I note the sensemaking elements (the set of activities) that constitute sensemaking. The sensemaking elements correspond to the *mediating processes* of my conjecture map - the set of interactions and artifacts produced by participants while engaging with the designed learning environment. My *high level conjecture* takes the form: "a set of task features which elicit the sensemaking elements increase the likelihood of students sensemaking". These task features then correspond to the *embodiment* component of my conjecture map - the material features which elicit the sensemaking elements of sensemaking. The arguments substantiating the embodiment, i.e. my *design conjectures*, have been detailed in Sections 4.3 to 4.6. Lastly, I note the theoretical

Table 4.1: *A brief summary of the task features entailing real-world context and mechanistic explanation, design conjectures associated with each feature, theoretical background for the design conjecture and the corresponding empirical evidence from the literature.*

Task-feature	Conjecture	Theoretical basis	Empirical evidence
Real-world context (RW)	If a task cues students (RW1) about the presence of a real-world context, then it is more likely to invoke students' everyday and curricular knowledge.	Semantic priming and Spreading activation theory	26;100;133
	(RW2) about the presence of a real-world context, then it is more likely to lead students to generate and connect diverse sets of ideas (conceptual, procedural and intuitive) from their knowledge system.	Search for Ideas in Associative Memory (SIAM)	137–142
Mechanistic Explanations (ME)	(ME1) cues students to generate mechanistic explanation(s), then it is more likely to invoke references to everyday and curricular knowledge. (ME2) to generate mechanistic explanation(s), then it is more likely to elicit mechanistic accounts.	Representational formats of mechanism knowledge	33;99;100;143;162
	(ME3) to generate mechanistic explanation(s), then it is more likely to cue generation and connection of diverse ideas from their knowledge system.	Theory of mechanistic reasoning	143;165;165;166;166

Table 4.2: *A brief summary of the task features entailing multiple representations and physical interpretations, design conjectures associated with each feature, theoretical background for the design conjecture and the corresponding empirical evidence from the literature.*

Task-feature	Conjecture	Theoretical basis	Empirical evidence
	If a task cues students		
Multiple Representations (MR)	(MR1) cues students to engage with multiple representations, then it is more likely to cue generation and connection of ideas from their knowledge system.	Cognitive Theory of Multimedia Learning (CTML)	175–177
	(MR2) cues students to engage with multiple representations, then it is more likely to nudge them in seeking coherence between ideas		169;172
Physical Interpretation (PI)	(PI1) cues students to interpret mathematical expressions in light of physical systems, then it is likely to facilitate generation and connection of ideas.	Mapping account/ Interpretation/ Inferential Conception/ Inferential Function	90;92;182
	(PI2) cues students to interpret mathematical expressions in light of physical systems, then it is likely to nudge them in seeking coherence between ideas.		29;55;87

conjectures about the *outcomes* of engaging in the sensemaking process from the literature (Section 2.2.3). Figure 4.2 highlights the contextual operationalization of conjecture mapping to my study.

4.7.2 Identifying task features that increase the likelihood of students sensemaking in physics

I identify the task features which increase the likelihood of students sensemaking in physics (**RQ2**) by leveraging contemporary theoretical ideas from cognitive psychology, education, and philosophy of science. These features include tasks cuing students about: (i) the presence of real-world context(s), (ii) to unpack the underlying mechanism of a phenomenon, (iii) to engage with multiple representations, and (iv) to physically interpret mathematical expressions. The identified features complement the contemporary pedagogical efforts in supporting students in making sense of their surrounding world using curricular ideas.

Several studies in science education have examined students' reasoning while engaging with real-world contexts (my first task feature). In addition to developing context-based pedagogical materials[183–186], researchers have analyzed students' cognitive, meta-cognitive, and affective behaviors while engaging with such materials[126;142;187;187;188;188–190]. These studies have noted context-based problems to enhance students' situational interest[142], motivation[183;191–193], along with improving attitudes towards science learning[194]. My work adds increased chances of engaging in sensemaking to this growing list. Ogilivie[195] indeed notes context rich, open-ended problems to be fertile grounds for students to notice inconsistencies in their knowledge systems - a crucial feature of the sensemaking process.

Researchers have also explored students' engagement with multiple representations (my third task feature) in physics. Coordination between representations – referred to as "representational fluency" – has been noted to assist students in invoking conceptual ideas not specified in the problem statement[169;196], leveraging information highlighted by representation(s)[176], and facilitating organization, prioritization, and communication of the contextual information[172;197–199]. My work adds to this list by noting that engaging with multiple representations leads to generating ideas and establishing coherence between the ideas thereby facilitating sensemaking.

Recent investigations have also explored the close association between sensemaking and modeling. Sirnoorkar *et al.*[100] note assembling of prior knowledge during sensemaking to entail construction of mental models about the target systems. The authors also note addressing and resolving the perceived inconsistencies during sensemaking to entail coherence seeking in the models, and testing them in light of their target systems. My identified task features complement these observations by facilitating promotion of sensemaking through modeling. Real-world systems (my first task feature) specify the nature of target systems, which when modeled, can increase the likelihood of students sensemaking in physics. Similarly, coordinating multiple representations, and physically interpreting mathematical results (the last two task features) specify the ways of establishing coherence and testing the merit of the models.

Chapter 5

Analyzing students' responses to tasks embodying features that promote sensemaking

In the previous chapter, I presented a theoretical argument in favor of four task features that increase the likelihood of students' sensemaking in physics. I argued that to promote sensemaking, tasks should cue students (i) about the presence of a real-world context; (ii) to generate a mechanistic explanation concerning the given context; (iii) to coordinate with multiple representations and (iv) to physically interpret mathematical expressions. I elucidated the design criteria through a conjecture map – a framework from design-based research – highlighting how each task feature attends to a characteristic feature of the sensemaking process. Figure 4.2 represents the conjecture map which relates the task-features ("Embodiment") and the characteristics of the sensemaking ("Mediating processes").

In this chapter I present

1. an example task embodying the proposed features.
2. empirical proof of the design conjectures by qualitatively analyzing a case study of a participant sensemaking on a physics problem.

3. proof-of-concept of the likelihood argument in eliciting sensemaking, i.e., how a task though embodying the proposed task features may not necessarily nudge students into sensemaking.

I present two case studies of problem solving and qualitatively analyze them through the lens of sensemaking and mechanistic reasoning. In the next section, I detail the theory of mechanistic reasoning and present the analytic framework in capturing the degree of mechanistic reasoning in our participants' work.

5.1 Theory: Mechanistic reasoning

Reasoning about the underlying mechanism of a phenomenon (the second cognitive element in our list) is a key component of the sensemaking process. As noted in the previous section, mechanistic reasoning is a form of causal reasoning that entails description of the events and the behavioral factors that lead to a phenomenon[96;143;163;164]. This form of reasoning entails generating explanations by transitioning from observable features of the phenomenon at the macro level to the underlying entities/processes (often at the micro level)[96;143]. The process of ascertaining the mechanism can further involve transitioning back from the micro to the macro features, and testing the validity of the generated explanation by varying the spatial or temporal organization of the entities/processes. This navigation across scalar levels – between observable features and underlying entities/processes, makes mechanistic reasoning a key component of the sensemaking process[96].

As an example, consider a common scenario frequently referred in introductory physics courses concerning the variation of a person's weight in an uniformly accelerating elevator. In order to understand the mechanism responsible for changes in the weight, one needs to identify the interactions of the forces on the rider's center of mass. The forces, as highlighted in Figure 5.1, include the rider's net force (ma), the person's weight (mg), and the normal force (N) exerted by the elevator's floor, where m, g, and a take usual notations for the rider's mass, the acceleration due to gravity, and the net acceleration of the elevator. Solving for the normal force during the elevator's upward and downward motion, one obtains:

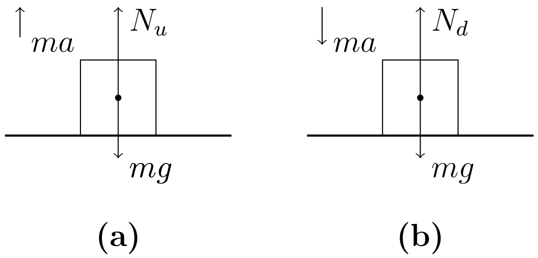


Figure 5.1: *Force diagrams describing the forces acting on a rider inside an accelerating elevator. The black dot represents the rider's center of mass. While (a) describes the scenario of elevator's upward motion, (b) describes the downward motion*

$$N_u = m\ (g + a) \tag{5.1}$$

$$N_d = m\ (g - a) \tag{5.2}$$

where N_u and N_d are the magnitudes of the normal forces experienced by the rider during the elevator's upward and downward motion. Since the apparent weight is the normal force experienced by the person, one can conclude from the above equations that the rider's weight while moving upward would be greater as compared to while moving downward. Furthermore, if the elevator were to be at rest (a=0), the normal forces in both the cases would be equal to the rider's weight (mg), which further validates our result.

A mechanistic account, such as the one observed in the above example involves making a transition from the observable variation of weight to the non-observable realm of forces' interactions on the rider. The validation of the generated explanation further involves transitioning back from the domain of forces to the observable behavior of the rider-elevator system under limiting conditions (i.e., with a being equal to zero).

5.1.1 Analytic framework on mechanistic reasoning

Russ *et al.*[143] put forward an analytic framework to capture the complexity of mechanistic reasoning in students' work. We adopt this framework to analyze the introductory student's mechanistic reasoning in our case studies (Section 5.3). The framework identifies the following nine hierarchical codes in the increasing order of their sophistication that relatively evidence mechanistic reasoning.

1. *Describing the target phenomenon.* Articulation or demonstration of a stable phenomenon concerning which reasoning about the underlying mechanism occurs. In our example, an identification of the variation in the apparent weight of the rider inside an elevator can evidence this code.

2. *Identifying set up conditions.* Noting the enabling conditions of the environment that describe spatial and/or temporal organization of entities that trigger regular changes of mechanism leading to the phenomenon. In the elevator scenario, the set up conditions can be the starting of the elevator motion (either upwards or downwards) that leads to the interaction of the forces on the rider can.

3. *Identifying entities.* The identification of objects or entities (such as floor of the elevator, rider's center of mass, and the involved forces) that play crucial roles in the manifestation of the phenomenon (for changes in the rider's weight inside an elevator).

4. *Identifying activities.* The identification of actions and interactions in which the entities engage in a mechanism correspond to this code. In our case, the motion of the elevator, directionality, and interactions of the forces correspond to the identification of the activities.

5. *Identifying properties of entities.* Identification or articulation of general properties of entities such as mass of the rider, and the nature of forces (gravity being directed towards the earth, normal force being perpendicular to the surface, etc.) that are instrumental in the mechanism.

6. *Identifying organization of entities.* The description of spatial organization, location, and the structure of entities, such as the rider being in contact with the elevator's floor, and the action of the forces on the riders' centre of mass corresponds to this code.

7. *Chaining: backward and forward.* Reasoning about the events that were crucial for the current behavior of systems (backward chaining), or that about the potential behavior of systems given its current state (forward chaining) mark this code. In our case, this code may correspond to analyzing the variation of the rider's weight upon stopping of the elevator.

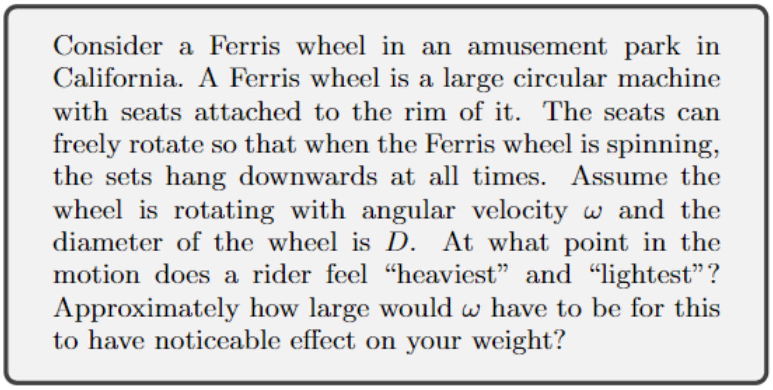
The authors further describe *Analogies* and *Animated models* as two additional codes which can accompany in students' mechanistic reasoning. Mapping of or comparing a known mechanism onto the phenomenon under consideration as a framework for understanding the given situation marks the *Analogies* code. When students employ external representations for illustrating any of the above-mentioned seven categories, it corresponds to the code of *Animated models.* The authors further notes higher the hierarchical codes, more the evidence of mechanistic reasoning in students' work.

We adopt this framework to analyze an introductory student's mechanistic reasoning on two physics problems which vary in terms of our proposed task features on sensemaking (Refer Sections 5.3 and 5.4).

5.2 Methodology

5.2.1 Data collection

Data for the current study is derived from two sets of think-aloud interviews which were conducted during spring of 2018 to investigate students' engagement with scientific practices[46]. Before going into further details, note that this study does not investigate students' reasoning about scientific practices for which the data were originally collected. However, the arguments made in the current work stem from observations of students' sensemaking in these interviews. Each of these interviews involved ten introductory students from a Midwestern US university solving a set of physics problems. The problems sets in both the interviews were designed using the Three Dimensional Learning Assessment Protocol (3D-LAP)[108]. While the first problem set consisted of seven tasks all of which elicited the



Consider a Ferris wheel in an amusement park in California. A Ferris wheel is a large circular machine with seats attached to the rim of it. The seats can freely rotate so that when the Ferris wheel is spinning, the sets hang downwards at all times. Assume the wheel is rotating with angular velocity ω and the diameter of the wheel is D. At what point in the motion does a rider feel "heaviest" and "lightest"? Approximately how large would ω have to be for this to have noticeable effect on your weight?

Figure 5.2: *Statement of the Ferris wheel task.*

scientific practice of "Using Mathematics", the second set consisted of nine tasks, of which three (tasks 7, 8 and 9) aligned with the scientific practice of "Developing and Using Models".

Both interview protocols involved asking participants to treat the problem-solving exercise as an untimed exam and that their participation had no bearing on their academic grades or courses. The participants were asked to express their thoughts out loud and the interviewer interacted with participants only when asked or to prompt them to express their thoughts aloud with questions such as '*What are you thinking?'* The participants were allowed to use calculators and were provided an equation sheet. Students were compensated with $20 for their participation.

Of the 320 instances of problem solving (16 problems attempted by 20 students) across the two sets of interviews, we present Luke's and Catherine's attempt at the "Ferris Wheel task" 5.2. Our choice to focus on Luke's and Catherine's attempts is guided by the two contrasting approaches demonstrated by the participants to the same task. While Luke made sense of the task by leveraging intuitive reasoning and multiple representations, Catherine reasoned about the same by manipulating a single equation. These two case studies demonstrated the feasibility of our *likelihood* argument in discussing about the task features. Additionally, the selection of the Luke's and Catherine's attempts were guided by their explicit and clear articulation of reasoning along with the audio/video clarity of the interviews.

Table 5.1: *Summary of the features of the Ferris wheel task reflecting the design task features of sensemaking.*

Proposed task features	Features of the Ferris wheel task
Real-world context	Ferris wheel as an amusement ride
Mechanistic explanation	Unpacking the underlying mechanism concerning the changes in rider's apparent weight inside the Ferris wheel.
Multiple representations	Construction of force diagrams and use of equations.
Physical interpretations	Interpreting force equations.

5.2.2 Context

Ferris-wheel task

The Ferris wheel task involves an amusement park ride in which a circular machine of diameter D with seats attached to its rim rotates vertically at angular velocity ω. The participants are asked to determine the points at which the riders would feel heaviest and lightest during the ride. One can determine this by considering the points on the ride at which the forces acting on the rider are either parallel or anti-parallel to each other as represented by points A and B in Figure 5.3. The figure illustrates the direction of the normal (N), gravitational (W) and centripetal (F_c) forces acting on the rider at these points. Since the riders' apparent weight is proportional to the normal force experienced by them, solving for the force at these points, one obtains:

$$N_A = m(g - a_c) \tag{5.3}$$

$$N_B = m(g + a_c) \tag{5.4}$$

where N_A and N_B are the normal forces experienced by a rider of mass m at points A and B respectively. g and a_c respectively represent the magnitudes of the acceleration due to gravity and centripetal acceleration. From the two equations, one can conclude that a

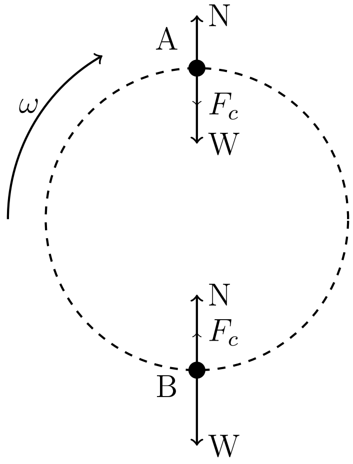


Figure 5.3: *Diagram highlighting the directions of weight and normal force on the rider at points A and B on the Ferris wheel. The black dot represents the rider's center of mass.*

rider feels lightest at the top and feels heaviest at the bottom.

5.2.3 Data analysis

Upon narrowing down on the case studies from the data set, I transcribed the interviews by considering the participants' verbal arguments, physical gestures, and their written solutions. The transcript was then arbitrarily segmented into three episodes, mainly to facilitate the ease of analysis. Explicit emphasis was laid on identifying participants noticing gaps or inconsistencies in their understanding. I looked for moments of students getting "stuck" as an evidence for the same. These instances were often accompanied by recurring pauses, shift in students' arguments, etc.[33;100].

5.3 Luke's attempt at the Ferris wheel task

In the rest of this section, I present Luke's approach to the Ferris wheel task split into three episodes with each episode simultaneously analyzed through the lenses of sensemaking and mechanistic reasoning. To guide the reader, we present below a brief summary of the participant's attempt.

Luke begins by going through the problem statement and decides to revisit the task after attempting remaining problems from the set. Upon revisiting, he ascertains the directions of the forces acting on the rider at various points on the Ferris wheel (Figure 5.4). He then notices an inconsistency in his understanding about the points on the Ferris wheel at which the riders would feel the maximum and the minimum normal force. In the end, he concludes that riders would feel the maximum (and minimum) normal force when the direction of the riders' acceleration is parallel (and anti-parallel) to the direction of gravity.

5.3.1 Episode 1

Luke reads the problem statement and constructs a representation highlighting the given diameter (D) and angular velocity (ω) of the Ferris wheel (Figure 5.4). He then states the direction of the wheel's acceleration as being tangential to its path and the centripetal acceleration being directed "inwards". After a careful re-read of the problem statement, he decides to move on to remaining problems from the set and mentions of revisiting the task later.

> [Reads the problem statement for 33 seconds.] *So, you have your Ferris Wheel here* [draws a circle] *and it has diameter of D and angular velocity of omega.* [Indicates D and ω in the diagram, refers to the equation sheet and the problem statement for 50 seconds] *So, acceleration of the wheel is always tangent to the path its traveling in.* [Looks away from the solution and the equation sheets] *Am I right about this? No, the centripetal acceleration is inward. I never liked these. I never liked these problems* [laughs]. *Umm, okay.* [Stares at the problem statement for 20 seconds] *I will come back to this later. Don't want to spend too much time on this now.* [Turns the page over to the next task].

Upon solving the remaining problems from the set, Luke revisits the Ferris wheel task and goes through the problem statement again. Initiating his second attempt at the task, he reasons about the forces acting on the rider by denoting the direction of the gravitational and centripetal forces at various points (A, B, C and D) in his constructed representation (Figure 5.4). He then looks for the centripetal force equation in the provided sheet.

> *Okay. Alright.* [Reads the statement again for 38 seconds] *At this point here, the cart is hanging down. And you have the force of gravity pointing downwards* [draws a box, draws and denotes a downward arrow with F_g at point C], *and you have the force, the centripetal force, right, because its, yeah* [pointing at the equation sheet and simultaneously looking at the problem statement]. *Am looking at circular motion* [Looking at equation sheet] *Which is pointing, which is inward. Centripetal force is inward [draws a downward arrow and indicates F at point C]. At your point here* [...] [indicates the direction of gravity, centripetal force at points A, B and D.]
>
> *So, in all of these instances, so the centripetal force is the mass times the centripetal acceleration. Is that right? It looks like its v squared over r. Is it, no? What is that?*

Referring to the centripetal force equation, Luke notes that all three quantities – rider's mass, radius and angular velocity of the wheel – tend to be constant in the given scenario. Consequently, he concludes that the direction of the centripetal force would determine the apparent weight of the rider.

> *And then, that's equal to mass times radius, times the rotational or the angular velocity. So, we know that the mass in all these situations is gonna be the same. So, really all that matters is the radius and the angular velocity. Yeah, angular velocity. But the radius is the same in all situations as well. So, only thing that's gonna matter is the angular velocity. Well, proportional to angular velocity, squared.*
>
> *Its also constant for the whole thing. So, here's what that matters is the direction of the force then, if we wanna know when the rider feels the heaviest and lightest.*

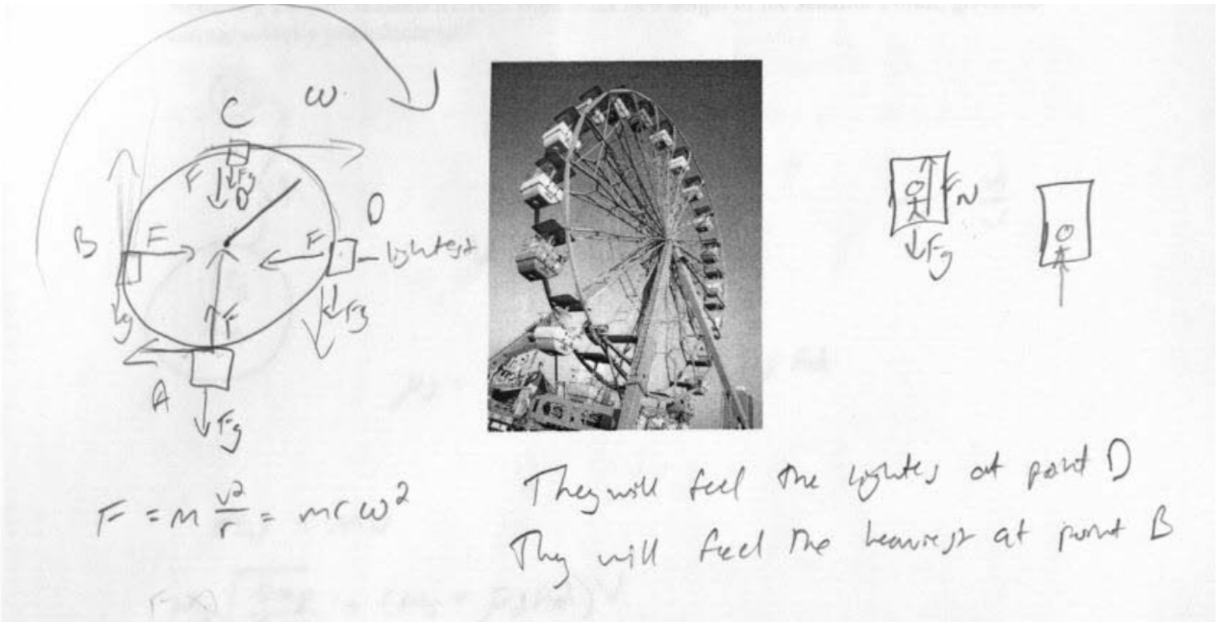


Figure 5.4: *Luke's written solution to the Ferris Wheel problem.*

Analysis: Luke's reasoning till this point reflects what Odden and Russ term as "assembling of prior knowledge" during sensemaking[40]. Luke primes his prior knowledge about the task by constructing a circular force diagram (on the left of Figure 5.4) and ascertaining the direction of the rides' gravitational and centripetal forces. Vignal and Wilcox note that such unprompted diagrams facilitate students in organizing and prioritizing the contextual information[200;201]. In addition, Luke physically interprets the centripetal force equation by noting all the involved terms – riders' mass, radius and angular velocity of the wheel – to be constant in the given context.

From the mechanistic reasoning angle, this episode reflects four codes of the analytic framework. Luke identifies *the target phenomenon*, by focusing on the changes in the rider's apparent weight in the Ferris wheel system ("*So you have your Ferris wheel here.*" and "*So, here's what that matters...*"). Thereafter, he identifies the *setup conditions* by emphasizing on the hanging cart and *identifies relevant entities* such as gravity, centripetal force, rider's mass, along with radius and angular velocity of the wheel. He further notes the *properties of entities* such as rider's mass, wheel's radius and its angular velocity to be constant across various points on the Ferris wheel.

5.3.2 Episode 2: Notices an inconsistency in understanding

Continuing further, Luke constructs additional pair of representations depicting the interplay between the gravitational and normal forces acting on the rider inside a cart (On the right of Figure 5.4). He then states that riders would feel the heaviest when they experience the maximum normal force and correspondingly feel lightest when they experience minimal normal force. Noting the gravity to remain constant at every point on the Ferris wheel, he concludes that the varying normal force determines the riders' apparent weight. He however, expresses a sense of puzzlement with the question "*Where will that be?*" expressing uncertainty in his understanding over the points at which the normal force would vary the most.

> *Well, we need to look at, inside the little carts you have the person and inside of every single cart, no matter where they are at, we have the force of gravity pointing down the person, and the normal force pointing upward from the floor, floor of the ride* [draws the figures on the right of the given Ferris wheel diagram].
>
> *So, they feel the heaviest, they feel the heaviest when the most force is being applied upward on them* [pointing his hand upwards] *and they feel lightest when they feel less force being pushed or pulled upward* [points his finger downward and then upwards]. *Because the force of gravity is gonna stay the same. The only thing that's changing is the normal force. Where will that be? So,* [looks at his circular diagram for ten seconds]

Analysis: This segment of Luke's transcript reflects noticing of a gap in his understanding about the points on the Ferris wheel at which the normal force on the rider would vary the most. This noticing is evidenced by the explicit puzzling question "*Where will that be?*" articulated upon concluding the variable nature of the normal force. This segment also reflects Luke's intuitive reasoning that more the upward force experienced by the riders, more would be their apparent weight ("*[...] they feel the heaviest when the most force is being applied upward on them*"). Similarly, he also notes that lesser the upward force on the riders, lesser would be the apparent weight (*"and they feel lightest when they feel less force being pushed or pulled upward"*). Finally, this segment also involves Luke revising his argument from the directionality of the centripetal force to the magnitude of the normal force as influencing the riders' apparent weight. Such revisions in arguments have been noted to be a key characteristic of the sensemaking process[21;40].

From the mechanistic reasoning perspective, this segment highlights Luke's continued efforts in *identifying relevant entities*. He identifies the varying normal force as yet another entity influencing the riders' apparent weight.

5.3.3 Episode 3: Addresses inconsistency

Accompanied by hand gestures, Luke intuitively argues that riders would feel the lightest and thus would experience the least normal force when their direction of motion is along the direction of the gravity. He identifies point D in his constructed circular representation as the point at which the riders would feel the lightest.

> *Well, at this point here* [pointing at D] *is where they will feel the lightest. Because, if we are assuming that it is rotating this way,* [draws the arc in the clockwise direction] *as you are moving downward* [gestures his hand downward], *the car is also moving, I don't know how to put this into numbers or equations or something to describe, but the cart will be moving downward and the person will be moving downward. So, the force between the person and floor is gonna be less because the vehicle, like the cart they are standing in, is also moving downward. So, there will be less of a force pushing on that person. They will be accelerating in the same direction. Yeah. So they feel the lightest, they feel the lightest at this point, I don't know how to, I guess we'll call this point A, this point B, this point C, and this point D. So,* [writes the concluding statement] *they will feel the lightest at point D.*

Extending the same argument, he identifies point B as the point at which the riders would feel the heaviest as their motion would be opposite to the direction of the gravity.

> *And they will feel the heaviest at point B. Because that's when the acceleration of the cart is upward, but you are accelerating downward. And I guess in this situation the cart will be accelerating that way, in this situation, they will will be accelerating towards left* [gestures hand leftwards], *in this situation the car will be accelerating downward.*

He then confidently sums up his arguments by reiterating his reasoning about the apparent weight experienced by the riders of the Ferris wheel.

> *So, you feel lightest when you will be accelerating in the same direction the person and the cart,* [gestures his fingers downwards] *and the person will be heaviest when the car is accelerating upward but the gravity is pulling you downward.*

Analysis: Luke's reasoning in this episode reflects his generation of explanations in response to the perceived gap in his understanding about the points at which riders would experience heaviest and lightest inside a Ferris wheel. He argues that riders would experience minimum normal force ("*So the force between the person and the floor is gonna be less..*") and thus feel lightest at point D as their direction of motion would align with the direction of the gravity. Along the same lines, he argues that riders would experience the maximum normal force and thus feel heaviest at point B at which when their direction of motion would be opposite to that of gravity ("*Because that's when the acceleration of the car is upward...*"). He identifies the points D and B as he explicitly acknowledges of not knowing how to expressing this argument through "*numbers of equations*". Luke's confident and clear articulation of the conditions in which a rider would feel the lightest and heaviest mark his resolution of the perceived gap and consequently the end of his sensemaking.

On the mechanistic reasoning front, this segment marks Luke identifying *setup conditions, activities*, and *organization of entities*, along with *chaining* codes in Luke's reasoning. He notes that when the Ferris wheel rotates (*setup condition*), the rider and the cart (*organization of entities*) too would be accelerating along the same directions (*identifying activities*). Similarly, he *chains* the same argument forward and notes that when moving upward at point B, riders would feel heaviest as the cart's acceleration would be directed opposite to that of the gravity. By virtue of the presence of all the codes, we note Luke's reasoning to be "mechanistically sophisticated"[143].

5.4 Catherine's attempt at the Ferris wheel task

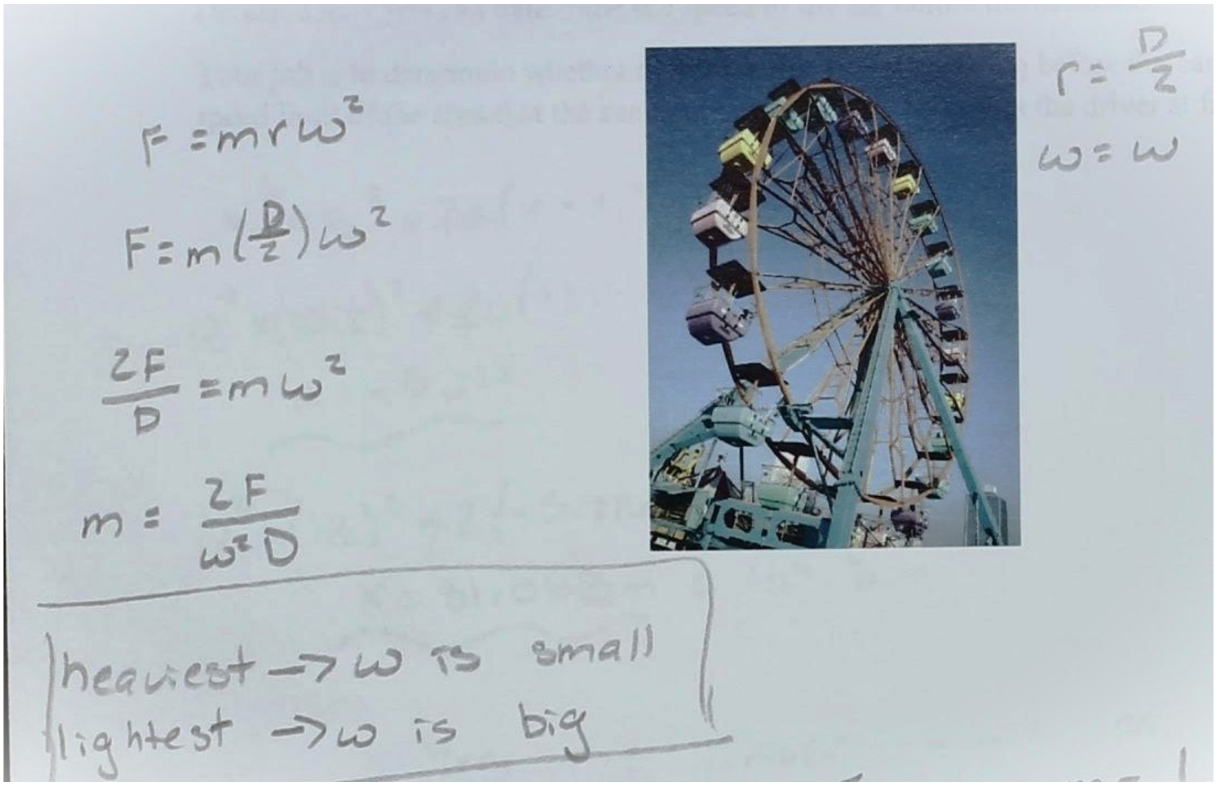
We now turn to another participant named Catherine, reasoning about the Ferris wheel task. Unlike the previous case study, Catherine's approach reflects the "plug and chug" approach in determining the points on the Ferris wheel where the riders would feel the heaviest and the lightest. I categorize the participant's transcript into three episodes and capture the mechanistic reasoning codes across the episodes.

5.4.1 Episode 1: Chooses centripetal force equation

Catherine initiates her approach by going through the problem statement and noting the "given" parameters. Noting the radius to be equal to half the diameter ("*D over 2*"), she narrows down on the centripetal force equation acknowledging the equation to entail all the given parameters from the problem statement.

> [Reads problem statement]. *So* [looks at equation sheet] *since we know that, we have the angular velocity and that the diameter of the wheel is "D", looking through the equation sheet* [looks at equation sheet] *there is only one equation that includes the angular velocity and then the radius of the circle. So since I have the diameter, I know that the radius* [starts writing] *will be equal to "D" over 2, and then the angular velocity will just be, the exact same. So using the centripetal force, we have* [reading from equation sheet, writing] *F equals "m" "r" omega squared. So we want to figure out when you'll feel the lightest and when you will feel the heaviest.*

She then substitutes radius as "D/2" in the centripetal force equation and seeks to solve for mass.



$r = \frac{D}{2}$

$\omega = \omega$

$F = mr\omega^2$

$F = m\left(\frac{D}{2}\right)\omega^2$

$\frac{2F}{D} = m\omega^2$

$m = \frac{2F}{\omega^2 D}$

heaviest → ω is small

lightest → ω is big

Figure 5.5: *Catherine's written solution to the Ferris Wheel problem.*

> *So* [writing] *I am going to sub in for "R" that way I have the problem in variables that they want in the problem.* [Looks back and forth between answer and equation sheets] *So, since we don't know what any of the actual like values are for this, we are going to just keep it in the variable form and we will solve for "m", the thing that we want to know when we feel the heaviest, and that correlates to your mass.*

Analysis: Catherine's reasoning in this episode reflects her efforts to coordinate the given information with the provided equations.

On the mechanistic reasoning front, her arguments in this episode reflect the evidence of two codes: "*describing the target phenomenon*" and "*identifying entities*". Catherine acknowledges the target phenomenon by noting: "*So we want to figure out when you'll feel the lightest and when you will feel the heaviest.*" She then identifies rider's mass, radius and angular velocity of the Ferris wheel as the relevant entities that play crucial role in the identified phenomenon.

5.4.2 Episode 2: Manipulates centripetal force equation

Continuing further, she manipulates the centripetal force equation and solves it for mass. The rearranged equation expresses mass in terms of the centripetal force, diameter and angular velocity of the Ferris wheel as shown in Figure 5.5. Noting the diameter to stay constant at all points on the Ferris wheel, she argues angular velocity in the equation's denominator to influence the rider's apparent weight.

> *So* [starts writing] *I will multiply it by the reciprocal of D over 2 which will get 2 F over D which is equal to 'm' angular velocity squared. And then divide by the angular velocity. M equal to F over angular velocity squared times D. From there we see that the angular velocity is squared which means that whenever we put in a negative value or a positive value it will come out positive... and so...* [reads paper for a few seconds] *so we can cross out the thought of having a negative angular velocity or a positive since it will be positive all of the time. The diameter of the Ferris wheel is the exact same no matter where you are it since it is a circle, so the biggest difference that would only impact your mass is the denominator of the fraction I have gets bigger. That would only be if the angular velocity gets a lot bigger or gets a lot smaller somehow.*

Analysis: This segment reflects Catherine's manipulation of the centripetal force equation and her interpretation of the reformulated equation.

On the mechanistic reasoning front, Catherine *Identifies properties of entities* by noting the constant nature of diameter and the variable nature of angular velocity.

5.4.3 Episode 3: Concludes

She interprets the formulated equation and concludes that higher angular velocity would make riders feel lightest, and lower angular velocity would make riders feel heaviest.

> *So you will feel* [starts writing] *the heaviest if the angular velocity is small just because it is in the numerator, so the numerator will have the bigger number. So when you divide you will get the bigger number for mass so you will feel lighter when the angular velocity is big.*

Analysis: This episode reflects Catherine's concluding claim by interpreting the centripetal force equation. Across the three episodes Catherine's reasoning reflects the "algorithmic" approach involving the manipulation of the centripetal force equation in light of the given information.

From the mechanistic reasoning perspective, this episode reflects her continued efforts to reason about the properties of involved entities, particularly the wheel's angular velocity. Collectively across the three episodes, Catherine's approach reflects three of the seven codes of mechanistic reasoning.

5.5 Discussion

We present two case studies involving participants – Luke and Catherine – solving the Ferris wheel task 5.2. The two case studies provide the proofs-of-concept of (i) how a specific feature in a task can elicit an element of sensemaking, (ii) the design conjectures and (iii) the likelihood argument on eliciting sensemaking. These arguments are detailed below.

5.5.1 Luke's approach as a proof-of-concept of design conjectures

In order to elucidate how a specific feature in a task can elicit an element of sensemaking, we below discuss each of the design conjectures highlighted in our conjecture map (Figure 4.2) manifest in the Luke's sensemaking. Below we present a detailed account of each design conjecture, and the corresponding evidence in the Luke's transcript. Each of the sensemaking elements highlighted in our conjecture map and their manifestation in Luke's transcript have been summarized in Table 5.2 and Figure 5.6. Furthermore, detailed arguments about

Table 5.2: *Summary of the evidence of sensemaking elements in Luke's transcript.*

Sensemaking elements	Evidence in Luke's transcript
Blending everyday and curricular knowledge	Intuitively argues that riders would experience minimal normal force (and thus feel lightest) at points where the direction of acceleration is parallel to the direction of gravity. Intuitively argues that riders would experience maximum normal force (and thus feel heaviest) at points where their direction of acceleration is opposite to the direction of gravity.
Mechanistic reasoning	Describing the target phenomenon: Explicit statement on changes in the rider's apparent weight inside the Ferris wheel. Identifying set up conditions: Rotation of Ferris wheel. Identifying entities: centripetal, gravitational, and normal forces, rider's mass, and, radius and angular velocity of the wheel. Identifying activities: acceleration of the riders along and opposite to the direction of gravity. Identifying properties of entities: rider's mass, wheel's radius and wheel's angular velocity to remain constant throughout its motion. Identifying organization of entities: position of cart and that of the person inside the wheel and directionality of the gravitational and normal forces on the rider. Chaining forward: directionality of the rider's acceleration and gravity at point B (as compared to point D).
Proposing and connecting ideas.	Centripetal acceleration is inward. Mass of the rider, radius and angular velocity of the wheel are all constant. Directionality of centripetal force influences riders' apparent weight. In every cart, gravity points down and normal force points upward. Magnitude of normal force influences riders' apparent weight. Heaviest when maximum normal force is experienced and lightest when minimum normal force is experienced. Directional alignment between riders' acceleration and gravity influences riders' apparent weight.
Seeking coherence between the ideas	Directionality of centripetal force influences riders' apparent weight. Magnitude of normal force influences riders' apparent weight. Directional alignment between riders' acceleration and gravity influences riders' apparent weight.

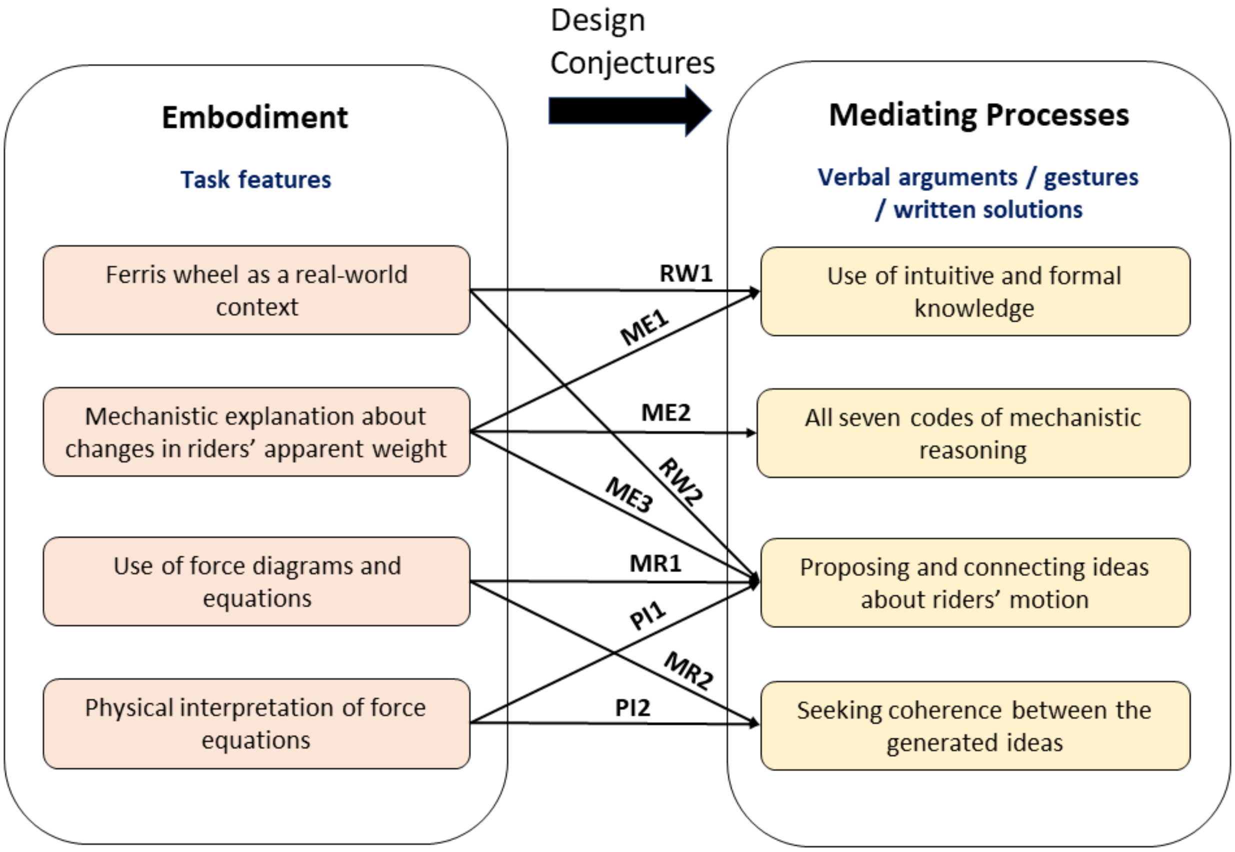


Figure 5.6: *Conjecture map highlighting the embodiment features of the Ferris wheel task and mediating processes/sensemaking elements evidenced in Luke's approach to the Ferris wheel task. Each of the mediating processes have been detailed in Table 5.2 and the design conjectures have been detailed in Section 5.5.*

how the Ferris wheel task embodies the proposed task features have been summarized in Section 6.3 and Table 5.1.

RW1

Conjecture: *If a task cues students about the presence of a real-world context, then it is more likely to invoke students' everyday and curricular knowledge.*

Evidence: Intuitively argues that

1. riders would experience minimal normal force (and thus feel lightest) at points where the direction of acceleration is parallel to the direction of gravity.
2. riders would experience maximum normal force (and thus feel heaviest) at points where their direction of acceleration is opposite to the direction of gravity.

RW2

Conjecture: *If a task cues students about the presence of a real-world context, then it is more likely to lead students to generate and connect diverse sets of ideas (conceptual, procedural and intuitive) from their knowledge system.*

Evidence: Argues

1. Centripetal acceleration of the cart's wheel to act inward.
2. Mass of the rider, radius and angular velocity of the wheel to be constant.
3. Direction of centripetal force to influence riders' apparent weight.
4. In every cart of the Ferris wheel, gravity points downward and normal force points upward.
5. Magnitude of the normal force to influence riders' apparent weight.
6. Riders feel the heaviest when they experience maximum normal force and lightest when they experience minimum normal force.

7. Directional alignment between riders' acceleration and that of gravity influences riders' apparent weight.

ME1

Conjecture: *If a task cues students to generate mechanistic explanation(s), then it is more likely to invoke references to everyday and curricular knowledge.*

Evidence: Intuitively argues that

1. riders would experience minimal normal force (and thus feel lightest) at points where the direction of acceleration is parallel to the direction of gravity.
2. riders would experience maximum normal force (and thus feel heaviest) at points where their direction of acceleration is opposite to the direction of gravity.

ME2

Conjecture: *If a task cues students to generate mechanistic explanation(s), then it is more likely to elicit mechanistic accounts.*

Evidence: All seven codes of mechanistic reasoning framework discussed in (Section 5.1) evidenced in Luke's reasoning. These are detailed below:

Describing the target phenomenon: Explicit statement on changes in the rider's apparent weight inside the Ferris wheel.

Identifying set up conditions: Noting rotation of Ferris wheel as the initiation of changes in riders' apparent weight.

Identifying entities: Noting centripetal, gravitational, and normal forces, rider's mass, and, radius and angular velocity of the wheel as entities influencing the changes in apparent weight.

Identifying activities: Noting riders' acceleration being parallel and anti-parallel to the direction of gravity.

Identifying properties of entities: Noting rider's mass, wheel's radius and wheel's angular velocity to remain constant throughout its motion.

Identifying organization of entities: Noting position of the cart and that of the person inside the wheel, and the directionality of gravitational and normal forces on the rider.

Chaining forward: Noting the directions of the rider's acceleration and that of gravity at point B (as compared to point D).

ME3

Conjecture: *If a task cues students to generate mechanistic explanation(s), then it is more likely to cue generation and connection of diverse ideas from their knowledge system.*

Evidence: Argues

1. Centripetal acceleration of the cart's wheel to act inward.
2. Mass of the rider, radius and angular velocity of the wheel to be constant.
3. Direction of centripetal force to influence riders' apparent weight.
4. In every cart of the Ferris wheel, gravity points downward and normal force points upward.
5. Magnitude of the normal force to influence riders' apparent weight.
6. Riders feel the heaviest when they experience maximum normal force and lightest when they experience minimum normal force.
7. Directional alignment between riders' acceleration and that of gravity influences riders' apparent weight.

MR1

Conjecture *If a task cues students to engage with multiple representations, then it is more likely to cue generation and connection of ideas from their knowledge system.*

Evidence: Arguing about

1. centripetal force equation:
   (a) Mass of the rider, radius and angular velocity of the wheel are all constant.
   (b) Direction of centripetal force influences riders' apparent weight.
2. force diagrams:
   (a) In every cart of the Ferris wheel, gravity points downward and normal force points upward.
   (b) Magnitude of the normal force influences riders' apparent weight.
   (c) Riders feel the heaviest when they experience maximum normal force and lightest when they experience minimum normal force.
   (d) Directional alignment between riders' acceleration and gravity influences riders' apparent weight.

MR2

Conjecture *If a task cues students to engage with multiple representations, then it is more likely to nudge them in seeking coherence between ideas.*

Evidence: Arguing about

1. centripetal force equation: Direction of centripetal force influences riders' apparent weight.
2. force diagrams:
   (a) Magnitude of normal force influences riders' apparent weight.
   (b) Directional alignment between riders' acceleration and gravity influences riders' apparent weight.

PI1

Conjecture: *If a task cues students to interpret mathematical expressions in light of physical systems, then it is likely to facilitate generation and connection of ideas.*

Evidence: Interprets centripetal force equation to conclude:

1. Mass of the rider, radius and angular velocity of the wheel are all constant.
2. Direction of centripetal force influences riders' apparent weight.

PI2

Conjecture: *If a task cues students to interpret mathematical expressions in light of physical systems, then it is likely to nudge them in seeking coherence between ideas.*

Evidence: Interprets centripetal force equation to conclude:

1. Direction of centripetal force influences riders' apparent weight.

5.5.2 Catherine's approach as a proof-of-concept of likelihood argument on eliciting sensemaking

This case study highlights the significance of an agentic approach in task design and the "*likelihood*" vocabulary in eliciting sensemaking. Catherine's approach reflects how factors beyond the task features such as personal epistemology[25] and in-the-moment framing of task's expectations[117] can influence students' reasoning about a task. Furthermore, the participant's attempt also substantiates the contextual nature of noticing inconsistencies in understanding.

Table 5.3 highlights the details of evidence of the sensemaking elements in Catherine's transcript. Figure 5.7 highlights the presence and absence of specific sensemaking elements in the participant's approach. I do not find evidence of the ME2, MR1, MR2, and PI2 design conjectures. The evidence design conjectures are detailed below.

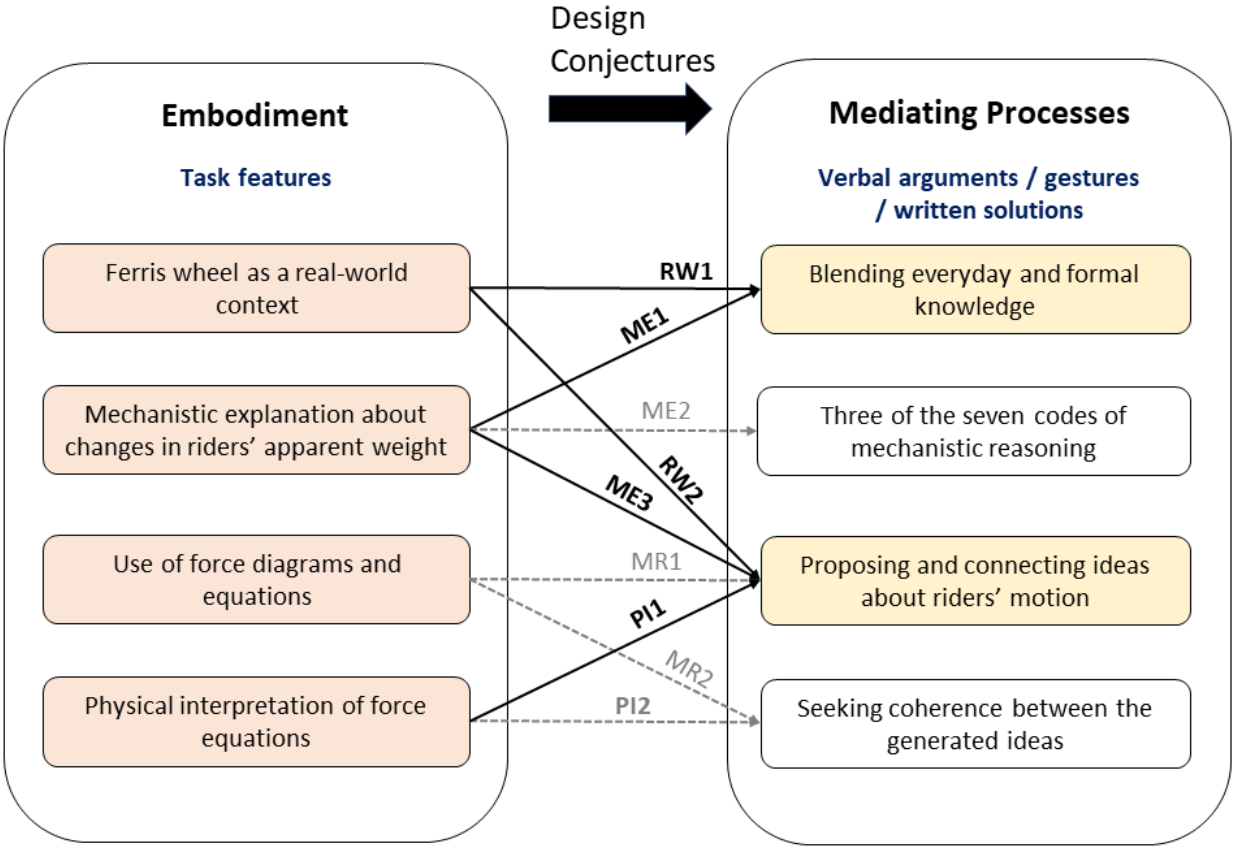


Figure 5.7: *Conjecture map highlighting the embodiment features of the Ferris wheel task and mediating processes/sensemaking elements evidenced in Catherine's approach to the Ferris wheel task. Each of the mediating processes have been detailed in Table 5.3 and the design conjectures have been detailed in Section 5.5.*

Table 5.3: *Summary of the evidence of sensemaking elements in Catherine's transcript.*

Sensemaking elements	Evidence in Catherine's transcript
Blending everyday and curricular knowledge	Solves the centripetal force equation for mass to reason about the riders' change in apparent weight. By virtue of ω being in the denominator of the simplified equation, riders would feel the heaviest for a smaller angular velocity. Similarly, riders would feel the lightest for a larger angular velocity.
Mechanistic reasoning	*Describing the target phenomenon:* Explicit statement on changes in the rider's apparent weight inside the Ferris wheel. *Identifying entities:* Identifies rider's mass, and, diameter and angular velocity of the wheel as factors influencing riders' apparent weight. *Identifying properties of entities:* wheel's diameter to remain constant and the angular velocity to vary during it's motion.
Proposing and connecting ideas.	Radius of the wheel as half the diameter. Angular velocity of the wheel to remain constant during its motion. Square of the angular velocity to always yield a positive value. Angular velocity of the wheel to vary during its motion.
Seeking coherence between the ideas	-

RW1

Conjecture: If a task cues students about the presence of a real-world context, then it is more likely to invoke students' everyday and curricular knowledge.

Evidence:

1. Solves the centripetal force equation for mass to reason about the riders' change in apparent weight.
2. By virtue of ω being in the denominator of the simplified equation, riders would feel the heaviest for a smaller angular velocity.
3. Similarly, riders would feel the lightest for a larger angular velocity.

RW2

Conjecture: If a task cues students about the presence of a real-world context, then it is more likely to lead students to generate and connect diverse sets of ideas (conceptual, procedural and intuitive) from their knowledge system.

Evidence: Argues

1. Radius of the wheel as half the diameter.
2. Angular velocity of the wheel to remain constant during its motion.
3. Square of the angular velocity to always yield a positive value.
4. Angular velocity of the wheel to vary during its motion.

ME1

Conjecture: If a task cues students cues students to generate mechanistic explanation(s), then it is more likely to invoke references to everyday and curricular knowledge.

Evidence:

1. Solves the centripetal force equation for mass to reason about the riders' change in apparent weight.
2. By virtue of ω being in the denominator of the simplified equation, riders would feel the heaviest for a smaller angular velocity.
3. Similarly, riders would feel the lightest for a larger angular velocity.

ME3

Conjecture: If a task cues students to generate mechanistic explanation(s), then it is more likely to cue generation and connection of diverse ideas from their knowledge system.

Evidence:

1. Radius of the wheel as half the diameter.

2. Angular velocity of the wheel to remain constant during its motion.

3. Square of the angular velocity to always yield a positive value.

4. Angular velocity of the wheel to vary during its motion.

PI1

Conjecture: If a task cues students to interpret mathematical expressions in light of physical systems, then it is likely to facilitate generation and connection of ideas.

Evidence: Interprets the centripetal force equation and argues that

1. Radius of the wheel to be equal to half the diameter.

2. Angular velocity of the wheel to remain constant during its motion.

3. Square of the angular velocity to always yield a positive value.

4. Angular velocity of the wheel to vary during its motion.

Catherine's approach to the Ferris wheel task (embodying the proposed task-features) presented evidence of the algorithmic mode of reasoning. Unlike Luke's case, Catherine's attempt did not reflect evidence of mechanistic reasoning, engagement with multiple representations, and the meta-cognitive activity of noticing gap in understanding.

Chapter 6

Sensemaking and Scientific Modeling: Intertwined processes analyzed in the context of physics problem solving

In this chapter, I examine the intertwining between the components of the sensemaking and modeling processes.

Students' meaningful engagement with curriculum has been a key area of focus in physics education research (PER). One thread of this endeavor focuses on students' use of knowledge from their lived experiences and from formal instruction to reason about their surrounding world. This form of reasoning, referred to as sensemaking[21], assists in generating new knowledge by leveraging one's existing ideas[79;91]. Sensemaking is also associated with the ways disciplinary experts engage in knowledge construction[23;24]. Given this significance, the field has experienced an increase in the investigations concerning the mechanisms[202] and the nature of tasks[28;29;33;113;203] associated with sensemaking. In addition to investigating the process, its significance, and ways of promoting it, researchers have also noted the close association sensemaking shares with modeling[21;27;35;36;38;39;46;73].

Modeling is a common cognitive process through which humans comprehend their surrounding world[35;36]. When modeling, the cognitive agent (the person engaging in modeling) abstracts and simplifies the target system to facilitate an explanation or a prediction[3;4;37;38]. During this process, ideas from one's knowledge are organized and applied to reason about various phenomena. Consequently, model-based reasoning can be a crucial component of sensemaking[27]. Several studies have noted that modeling and sensemaking share several common features including the common objective 'to figure it out'[21]. The two processes tend to consist of multiple phases including priming of prior knowledge of the concerned context, noticing discrepancies while reasoning, and generating new knowledge by connecting one's existing ideas[39;40].

Furthermore, the connection between sensemaking and modeling has been described quite explicitly. For instance, Sands[36] proposes the 'ACME protocol' elucidating mathematical modeling as the process of making sense of the physical world through three distinct stages. These include (i) Assessing the problem (making sense of the problem), (ii) Constructing the Model (making sure that the model makes sense), and (iii) Evaluating the model (making sense of the physical world using the model). These stages encompass the construction of a mental model as the modeler makes sense of the problem, expresses the ideas of the mental model mathematically (particularly during problem solving), and tests the model through its use to generate explanations or make predictions.

However, most of the above-mentioned accounts in the literature on the interplay between sensemaking and modeling focus on describing what modeling entails, but rarely articulate what constitutes making sense of something. An explicit description is important since sensemaking is a complex cognitive process in itself[21;41]. A better understanding of how engaging with models facilitates sensemaking requires a detailed account of the sensemaking process. Additionally, I are not aware of any demonstration of the association between the components of the two processes through explicit frameworks. A nuanced description of the association between the corresponding elements can further the current understanding about the two widely studied processes.

In the current work, I probe the association between the two processes by qualitatively analyzing the case studies of two students, Matthew and Ken, sensemaking about a physics problem by modeling the given context. I examine the students' modeling by noting their construction of mental models[42], and their subsequent expression of the model's ideas using Suárez's Denotative Function (DF), Demonstration (D), and Inferential Function (IF) – the DFDIF account of modeling[4]. Additionally, I also examine the students' sensemaking through Odden and Russ's four stages of the Sensemaking Epistemic Game[40]. While Matthew completes the task (albeit with minor deviations from expected calculations), Ken opts to quit the problem-solving exercise abruptly. In Matthew's case, I note his construction of a mental model, and engagement with the DFDIF components to entail navigation through the stages of the Sensemaking Epistemic Game. Observations from Ken's case suggest that when the barriers to modeling are experienced, sensemaking is inhibited.

6.1 Theoretical frameworks

The contemporary descriptions on the intertwining between modeling and sensemaking lack (i) a detailed account of sensemaking, and (ii) a framework-based methodological approach in unpacking the intertwining between the two processes. I address the first concern by discussing the sensemaking process through the lens of the Sensemaking Epistemic Game[40] which succinctly describes how the sensemaking process begins, proceeds, and terminates through four distinct stages. On the other hand, I adopt the DFDIF account[4], which describes the modeling process in terms of three components thereby facilitating a categorical analysis of the association between modeling and sensemaking (thus addressing the second concern).

Below, I describe the adopted frameworks on sensemaking and modeling in addition to noting their relation with Sands' "ACME protocol"'[36], which is the most explicit discussion about the intertwining between the two processes. The ACME protocol describes mathematical modeling through three stages: (i) Making sense of the problem - developing a qualitative mental model about the target system, (ii) Making sure that the model makes sense - translating the qualitative ideas into mathematical relationships, and (iii) Making sense of the physical world using the model - physically interpreting the established mathematical relationships.

Table 6.1: *A brief description of the stages in the Sensemaking Epistemic Game. Stages 1-3 correspond to sensemaking.*

Stages of Sensemaking Epistemic Game	Description
0. Assembling of a knowledge framework	Prior to sensemaking, students prime their existing knowledge on the concerned task.
1. Noticing a gap in knowledge	Students transition into sensemaking by noticing an inconsistency between existing knowledge and the knowledge required from the task.
2. Generating an explanation	Explanations are generated in response to the noticed inconsistency by connecting one's existing ideas.
3. Resolution	Students conclude the sensemaking process upon generating a satisfactory explanation.

6.1.1 Sensemaking Epistemic Game

In the current work, I adopt Odden and Russ's following account of sensemaking as:

> *a dynamic process of building or revising an explanation in order to 'figure something out' - to ascertain the mechanism underlying a phenomenon in order to resolve a gap or inconsistency in one's understanding. One builds this explanation out of a mix of everyday knowledge and formal knowledge by iteratively proposing and connecting up different ideas on the subject. One also simultaneously checks that those connections and ideas are coherent, both with one another and with other ideas in one's knowledge system*[21].

Odden and Russ also propose the Sensemaking Epistemic Game[40] framework to describe how the process begins, progresses, and concludes. This framework adopts the construct of epistemic games[81] to characterize the trajectory of sensemaking. Epistemic games correspond to the set of rules employed when undertaking a scientific inquiry towards the goal of the task called a *target epistemic form*. These epistemic forms vary with the task under investigation. When playing any game, I abide by a certain set of rules which are referred to as the *constraints* of the game.

Along with the target epistemic forms and constraints, epistemic games also have *entry conditions, moves, and exit conditions.* The circumstances around which a person begins to play the game are called *entry conditions.* These conditions are often triggered by the nature of the inquiry. *Moves* are the set of actions taken at various stages of the game, and *exit conditions* represent the circumstances in which the inquiry is terminated.

As an example, consider determining the displacement between the given points A and B in a two-dimensional Euclidean space. Let the points be defined by the coordinates A $(x = 2, y = 3)$ and B $(x = 5, y = 1)$. The *target epistemic form* of this inquiry (epistemic game) would correspond to ascertaining the magnitude and direction of the displacement vector defined between the points. The *constraint* in such a case would be to obtain a single numerical value along with its direction. The *entry and exit conditions*, that is, how one initiates and terminates the approach, can vary according to whether the prompt is open-ended or a multiple choice. Performing algebraic calculations on the given data or manually plotting the points on a graph would count as the *moves* of this game.

Odden and Russ adopt the epistemic games construct and describe the trajectory of sensemaking through the following four stages of the Sensemaking Epistemic Game[40]:

Stage 0: Assembling a knowledge framework

Before sensemaking, students prime their existing knowledge on the given domain. This can include recalling the general understanding of the task-related concepts, priming on the provided information, etc. This stage of assembling prior knowledge is highly dependent on the activity's context and forms the precursor to sensemaking.

Stage 1: Noticing a gap or inconsistency

Students notice a gap between their existing knowledge and the knowledge expected from the task. The noticing of the gap in one's knowledge system marks the entry condition into the epistemic game. This is often accompanied with recurring pauses, articulation of vexing questions, etc.

Stage 2 Generating an explanation:

In response to the noticed inconsistency, explanations are generated by connecting one's existing ideas and seeking coherence between them. These explanations are often drawn from the knowledge gained through one's lived experiences and formal instruction.

Stage 3 Resolution:

Students reach the target epistemic form (goal of the task) upon generating a satisfactory explanation which addresses the noticed inconsistency. Resolution of the inconsistency marks the exit condition from the Sensemaking Epistemic Game. This stage is highlighted by confident articulation of the explanation or a claim with appropriate justification.

Table 6.1 summarizes the four stages of the Sensemaking Epistemic Game. It is worth noting that sensemaking can occur despite inconclusively addressing or not at all addressing the perceived gap. That is, one can engage in sensemaking even upon quitting the sensemaking epistemic game abruptly (though upon noting the gap in understanding). One can also engage in sensemaking despite resolving the perceived gap through incorrect solutions. Thus, sensemaking can occur even if there is *incomplete resolution* or *incorrect resolution*.

While the sensemaking epistemic game framework is consistent with the adopted definition of sensemaking, it however differs from the Sands' 'ACME protocol'[36]. The difference mainly lies in the meta-cognitive feature of noticing discrepancies while reasoning. Unlike the ACME protocol, the objective of sensemaking (according the Sensemaking Epistemic Game) is to address a perceived discrepancy in one's knowledge system.

In the sections to follow, I analyze students' sensemaking on a physics problem by categorizing their approach in terms of the above-mentioned stages of the Sensemaking Epistemic Game. In order to observe the intertwining between the components of sensemaking with modeling process, the same problem solving approaches are analyzed through the components of modeling described below.

6.1.2 Modeling

Similar to sensemaking, the notion of engaging with models has been a fragmented construct in terms of its diverse characterization. Models have typically been regarded as representations standing completely or partially in isomorphic relation to their target systems[204;205]. However, recent discussions in the philosophy of science literature have increasingly emphasized the role of human agent's cognitive interests in modeling a system[44;206;207]. Consequently, in this work, I define modeling as

> *an activity engaged in by a cognitive agent involving abstraction and simplification of a phenomenon in order to generate explanations or predictions*[38;206].

In addition, consistent with the recent arguments in the literature, I consider modeling to involve an initial construction of a qualitative mental model, and subsequently expressing the model's ideas through external representations[36].

Mental models are *internal representations which act as structural analogue of situations or processes*[42] and have been noted as a precursor of constructing external representations[36]. Human beings' interaction with the physical world is mediated by mental models which are based on one's lived experiences and social interactions[208]. As representations of the physical world, the mental models are incomplete as they do not possess all the requisite information to comprehend and explain a complex context. Thus in the initial stages of a complex activity such as physics problem solving, mental modeling often involves making sense of the contextual information in light of one's existing knowledge[36;209].

The external representations, on the other hand, often reflect the qualitative ideas contained in the mental model, and depict the specific features of the target system to be modeled. Suarez[4], building on the work of Hughes[3], elucidates the process of modeling through external representations in terms of the following three components:

1. **Denotative Function (DF):**

Elements of the target system are specified by elements of the external representation when abstractly portraying a phenomenon. This feature of making a transition from the 'physical world' to the 'model world' marks this component of modeling.

2. **Demonstration (D):**

The elucidation of the internal dynamics between the designated elements in the representation marks the second component of modeling. In physics, this is typically achieved by establishing mathematical relationships between the denoted elements of the representation.

3. **Inferential Function (IF):**

The theoretical relationships obtained in the Demonstration stage are then interpreted in terms of the target system. The reverse-transition from the 'model world' to the physical world corresponds to the last component.

Hughes[3] and Suárez[4] illustrate the above-mentioned components by referring to Galileo's kinematical problem introduced in the Third Day of his Discourses Concerning Two New Sciences[2]. In Proposition I, Theorem I, Galileo demonstrates that two objects, one accelerating from rest and the other traveling at constant speed, would cover the same distance in a given time if the former's terminal speed is twice the latter's uniform speed. He elucidates this result by representing the kinematical scenario through a geometrical representation as shown in Figure 6.1.

In the figure, the length of the line segment AB represents time taken by the accelerating object to cover a given distance (say x units). AE is drawn such that increasing length of the horizontal lines from AB to AE represent the increasing degree of instantaneous speeds of the accelerating object. The parallelogram AGFB is constructed such that FG is parallel to AB and bisects EB. FG then represents the time taken by a uniformly moving object (as evidenced by the equidistant lines from AB) in traversing the distance x. Areas of the parallelogram AGFB and triangle AEB (i.e., the distances travelled by the two objects) can be shown to be equal provided the terminal speed of the accelerating object EB is twice the uniform speed FB of the other.

The above proof is a clear example of modeling in which a physics problem is expressed as a problem in geometry. In order to abstractly express the kinematic scenario as a geometrical representation, the physical quantities (speed and time) are denoted as horizontal and vertical line segments. This feature of establishing a relationship between physical quantities and elements of the representation constitutes the Denotative Function component of modeling.

Table 6.2: *A brief description of the DFDIF components of modeling.*

DFDIF Components of Modeling	Description
Denotative Function (DF)	Elements of the target system are specified by elements of the representation when abstractly portraying a phenomenon.
Demonstration (D)	The elucidation of the internal dynamics between the denoted elements.
Inferential Function (IF)	The mapping of theoretical conclusions onto the target system.

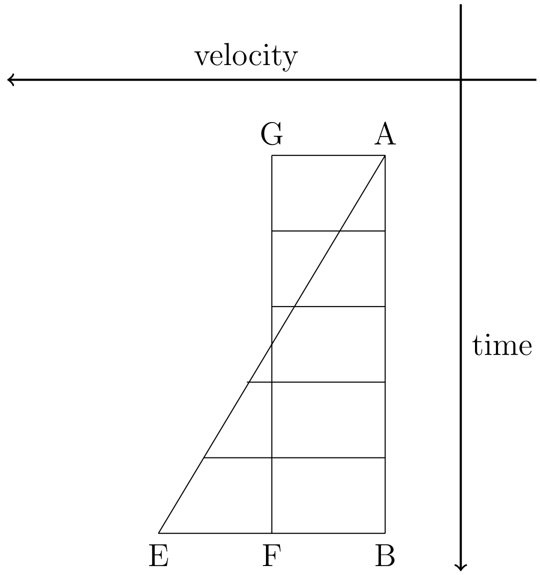


Figure 6.1: *Modified version of the Galileo's diagram relating the distances covered by a uniformly accelerating object starting from rest, to an object moving with constant speed. The length of vertical lines represent the time taken by the two objects to traverse a constant distance. The horizontal lines correspond to the increasing speeds. Refer*[2–4] *for the original diagram.*

Thereafter, areas of the parallelogram AGFB and triangle AEB are equated, with EB being twice EF. This elucidation of the internal dynamics between the denoted line segments marks the second component, i.e., Demonstration. Lastly, the geometrical results obtained in the Demonstration stage are interpreted in terms of kinematics reflecting the Inferential Function component.

the adopted DFDIF framework on modeling shares several common features with the Sands' 'ACME' protocol[36]. The first stage of the protocol (assessing the problem) broadly aligns with the Denotative Function component wherein elements of the target system are identified. The second and the third stages (constructing and evaluating the model) reflect the remaining two components of the DFDIF framework wherein the identified elements are related (Demonstration) in order to generate explanations/predictions (Inferential Function).

6.2 Research Questions

In the rest of this paper, I seek to address following research questions:

1. How does engagement with mental models and the DFDIF components associate with navigation through the stages of the Sensemaking Epistemic Game?
2. How does construction of mental models and engagement with DFDIF components influence students' sustained engagement in sensemaking?

Answering the above research questions will further the current understanding of the intertwining between the elements of the two processes. A nuanced understanding of this intertwining also can assist in sustaining and promoting sensemaking through modeling in our learning environments.

6.3 Methodology

6.3.1 Data Collection

I adopt case study analysis to unpack the intertwining of modeling components with the stages of the Sensemaking Epistemic Game. This approach allowed us to deeply analyze the intertwining between the two processes through an in-depth analysis of a target phenomenon[210]. The case studies are derived from a pool of video data consisting of students from a Midwestern US university participating in think-aloud interviews[70]. The interviews were comprised of individual students (N=10) articulating their thoughts aloud while solving a set of physics problems. The interview protocol involved asking participants to treat the problem-solving exercise as an exam and were informed that the exercise was untimed and their participation had no bearing on their academic grades. Students were compensated with $20 for their participation. The interviewer interacted with participants only when asked, or to prompt them to express their thoughts aloud with questions such as "*What are you thinking?*". The participants were allowed to use calculators and were provided with an equation sheet along with the problem set.

The problem set consisted of nine introductory college physics tasks designed to investigate students' engagement in Scientific Practices[46]. Of these nine problems, three (problems 7, 8 and 9) were specifically designed using the Three-Dimensional Learning Assessment Protocol (3D-LAP)[108] to elicit the Scientific Practice of 'Developing and Using Models'[46]. In the rest of this paper, I focus on students' responses to problem 7 - the 'Gravitron task' (Section 6.3.2). I focus on this particular task for two reasons. First, unlike the responses to the first six problems, a majority of students in my data reasoned about the Gravitron task by making arguments from their lived experiences and through physical gestures. This characteristic shift in approach in students' approaches caught my attention. Second, the Gravitron task is also the first 3D-LAP question in the problem set which students encountered and presumably engaged in the modeling practice.

Of the 10 participants, I focus on Matthew's and Ken's (both pseudonyms) approaches to the Gravitron problem. In addition to the audio/video clarity of their interviews, my choice to focus on the two students' approaches is informed by their clear articulation of arguments while solving the problem. The clarity of the arguments played a key role in identifying various elements of modeling and sensemaking processes. While both students sought to make sense of the Gravitron's scenario, they differed in their sustained engagement. Matthew's attempt involved successful navigation through the four stages of the epistemic game whereas Ken chose to quit the process abruptly. Comparison between these two case studies offer valuable insights on the circumstances in which students may opt to abandon sensemaking.

You are asked to design a Gravitron for the county fair, an amusement park ride where the rider enters a hollow cylinder, radius of 4.6 m, the rider leans against the wall and the room spins until it reaches angular velocity, at which point the floor lowers. The coefficient of static friction is 0.2. You need this ride to sustain mass between 25-160 kg to be able to ride safely and not slide off the wall. If the minimum ω is 3 rad/s, will anyone slide down and off the wall at these masses? Explain your reasoning using diagrams, equations and words.

Figure 6.2: *Statement of the Gravitron problem*

6.3.2 The Gravitron Problem

The Gravitron task (see Figure 6.2 for the problem statement) involves a rotating cylindrical amusement park ride in which the rider leans against the wall. With the given parameters, students are asked to determine whether riders would slide off the Gravitron's wall.

One of the ways to approach this task is by noting the forces acting on the rider through a force diagram as shown in Figure 6.3. If a rider is to be suspended without slipping, the vertically upward acting frictional force (F_f) offered by the wall must be at least equal to the downward pull of gravity (F_g). Mathematically, this argument can be expressed as:

$$F_f = F_g \tag{6.1}$$

Since the maximum friction offered by a surface is equal to the surface's coefficient of static friction (μ) times the normal force (N), and as the normal force provides the necessary centripetal force, the above equation can be simplified as

$$\mu(mr\omega^2) \geq mg \tag{6.2}$$

where m represents the rider's mass, r ($= 4.6m$) represents the Gravitron's radius, ω ($= 3rad/s$) represents the rider's angular velocity and g ($= 9.8m/s^2$) represents the acceleration due to gravity. The simplification of Equation 6.2 leads to the condition

$$\mu r\omega^2 \geq g \tag{6.3}$$

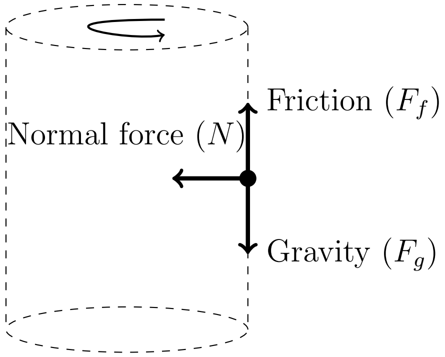


Figure 6.3: *A force diagram representing the gravity, friction, and normal force acting on a rider in the drum of a Gravitron amusement park ride (see task statement in Figure 6.2). The dot in the diagram represents the rider's center of mass.*

By substituting the given parameters, one observes that the above inequality does not hold true and consequently a rider would slide off the Gravitron's walls.

From the epistemic games perspective, the Gravitron task's *target epistemic form* corresponds to the final claim about the rider's status inside the Gravitron. This claim is further *constrained* by the two possibilities about the riders either falling off, or holding up against the Gravitron's walls.

6.3.3 Data analysis

After narrowing down the case studies, the first author iteratively viewed and transcribed Matthew's and Ken's interviews, taking into account the participants' speech, physical gestures, and written solutions. The transcript – documented description of participants' attempts in terms of verbal arguments, gestures and written solutions – was then examined and segmented into the stages of the Sensemaking Epistemic Game based on the discourse markers discussed in the literature[26;211]. As evidence for the Noticing of an Inconsistency stage, I looked for instances of students 'getting stuck', which were cued by markers such as articulation of puzzling questions in their arguments[26] accompanied by representational gesturing[211]. Instances preceding this stage where students recalled or gathered task-related information marked the Assembling of a Knowledge Framework stage of sensemaking. The Generation of an Explanation stage was identified by noting the mathematical or conceptual

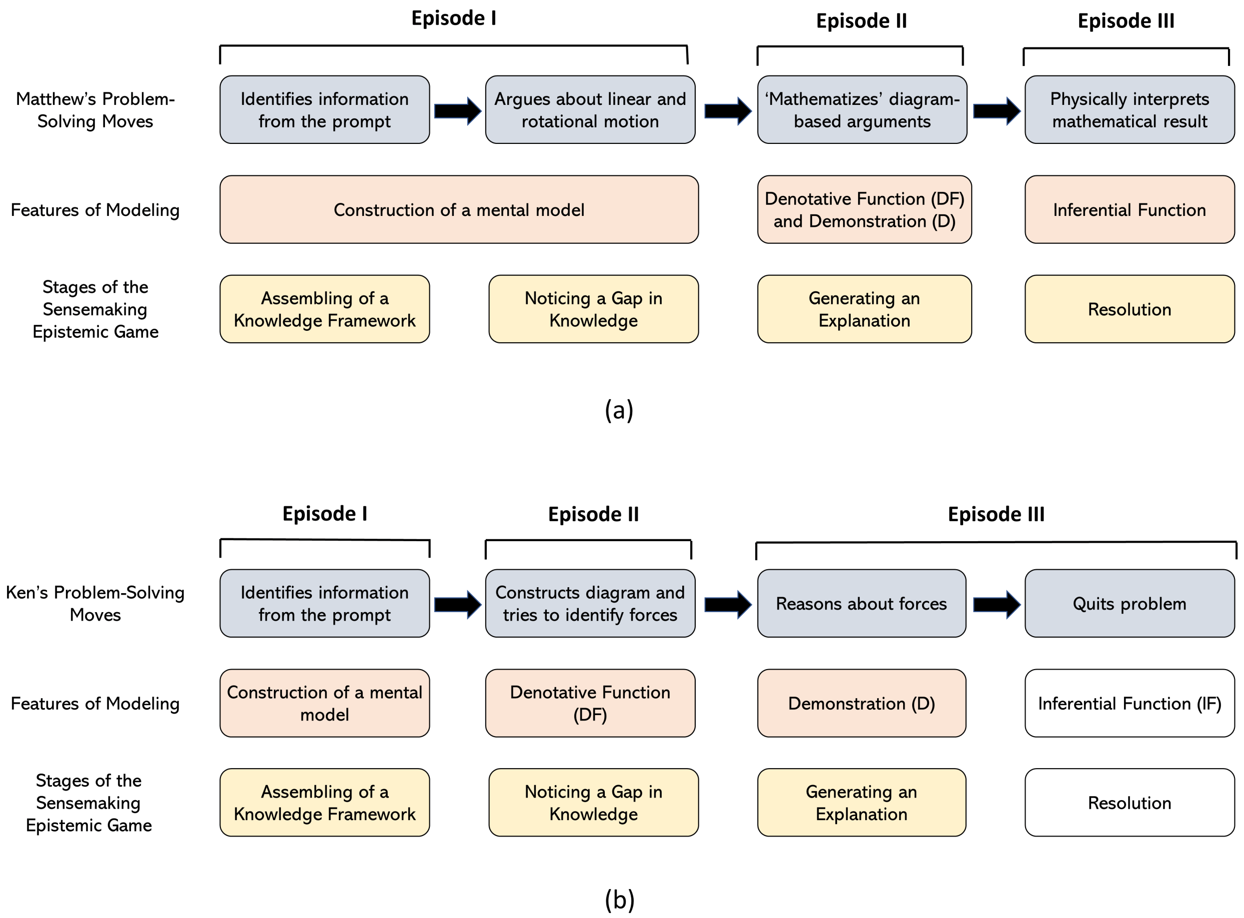


Figure 6.4: *Alignment of Matthew's (a) and Ken's (b) problem-solving moves with the stages of the Sensemaking Epistemic Game and the features of modeling across the three episodes. In both the figures, the top rows describe the participants' problem-solving moves in their attempt at the Gravitron task, the middle row highlights their construction of a mental model and their engagement with the DFDIF components of modeling, and the bottom row identifies the stages of the Sensemaking Epistemic Game evidenced during their attempts.*

arguments that the participants made after noticing inconsistencies in their understanding. Lastly, students' correct and/or complete articulation of their final claim marked the 'Resolution' stage of the epistemic game.

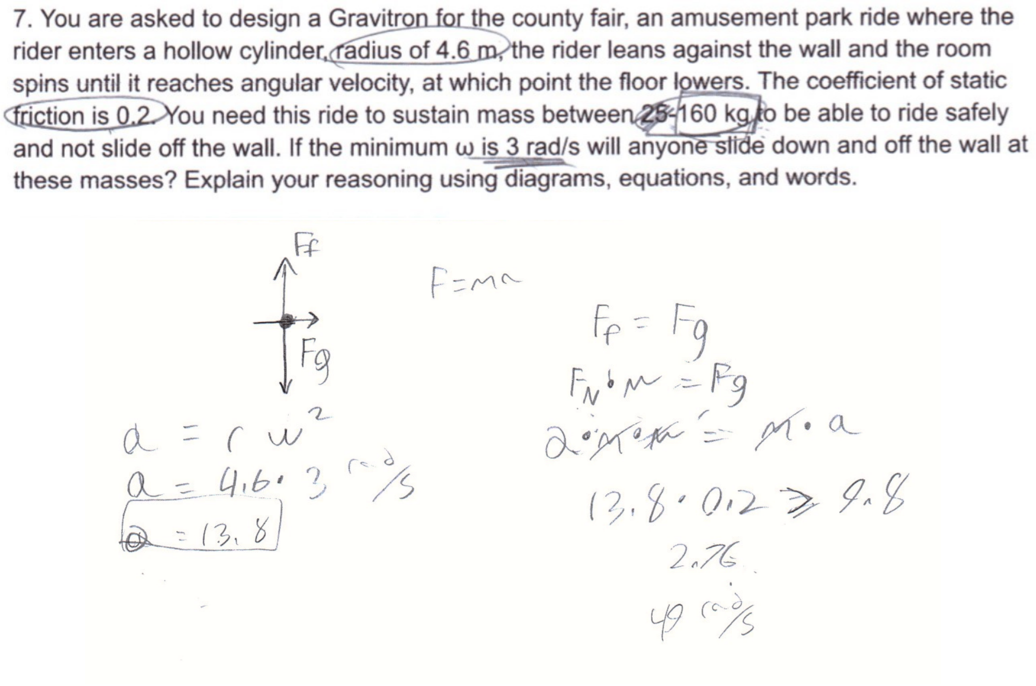
To capture the modeling component in the participants' work, the first author segmented the same transcript into two parts: before and after construction of an external representation. Students' arguments before construction of external representations, particularly those which accompanied representational gesturing[42;211], and those reflecting a sense of 'incompleteness or uncertainty'[45] cued about their engagement with mental models. Note that the objective of the current work is neither to characterize nor analyze the participants' mental models, but rather to identify the instances of their mental modeling to make comparisons to the Sensemaking Epistemic Game.

I identified the modeling's DFDIF components by noting the participants' arguments during and after construction of external representations (diagrams or equations). For instance, associating forces with arrows in a force diagram, or designating an equation's algebraic symbols with Gravitron's parameters indicated the Denotative Function in the students' work. Highlighting the directional relationships between the denoted arrows, or establishing mathematical relationships between algebraic symbols reflected the Demonstration component. The physical interpretation of the relationships established in the Demonstration stage marked the modeling's Inferential Function.

The segmented transcripts of the students' approaches in terms of stages of the epistemic game and components of modeling were then compared. The relative positioning of the modeling activities (mental modeling and engaging with DFDIF components) with the stages of the Sensemaking Epistemic Game highlighted the association between the two processes (Refer Figure 6.4). In the following two sections, I provide a detailed account of Matthew's and Ken's approaches, categorized into modeling activities (construction of mental model and engaging with DFDIF components) and the stages of the Sensemaking Epistemic Game.

6.4 Matthew's attempt at the Gravitron task

Matthew begins by noting the given information before noticing a gap in his understanding. He addresses the perceived gap by constructing a force diagram, and translating the diagram-based arguments into an algebraic inequality. The validity of the inequality makes Matthew to conclude that riders would slide down the Gravitron's walls under the given conditions.



7. You are asked to design a Gravitron for the county fair, an amusement park ride where the rider enters a hollow cylinder, radius of 4.6 m, the rider leans against the wall and the room spins until it reaches angular velocity, at which point the floor lowers. The coefficient of static friction is 0.2. You need this ride to sustain mass between 25-160 kg to be able to ride safely and not slide off the wall. If the minimum ω is 3 rad/s will anyone slide down and off the wall at these masses? Explain your reasoning using diagrams, equations, and words.

$F = ma$

$a = r\omega^2$

$a = 4.6 \cdot 3$ rad/s

$\boxed{a = 13.8}$

$F_f = F_g$

$F_N \cdot \mu = F_g$

$a \cdot \cancel{m} \cdot \mu = \cancel{m} \cdot a$

$13.8 \cdot 0.2 \geq 9.8$

2.76

49 rad/s

Figure 6.5: *Matthew's written solution to the Gravitron problem.*

6.4.1 Episode I

Matthew initiates by highlighting the parameters from the problem statement (as observed in Figure 6.5) and says:

> *So, I just went through and circled all the important information. Here it talks about the mass of the rider.*

He then refers to the ease of free fall by a heavier person before shifting his argument to angular momentum. This argumentative shift is accompanied by a hand gesture and intermittent pauses.

> *If the person who weighs more, then its gonna be easier for them to drop straight down* [gesturing his pencil downwards], *and harder for them to...* [looks at the problem sheet], *is that right?*
>
> [Pauses for 20 seconds while looking back and forth between the problem and the equation sheet.]

> *Since its angular momentum, they will have...* [draws and erases a diagram and reads the problem statement again. Pauses for 17 seconds]

Matthew's construction of a qualitative mental model

The transcript till this point reflects Matthew's construction of a qualitative mental model. Matthew makes an incomplete reference to the rider's free fall ("[...]*harder for them to...*") by abruptly interrupting his statement with the question "*is that right?*" After a pause of 20 seconds, he refers to the rider's rotational motion through the statement "*Since its angular momentum...*". This shift in arguments, from the rider's linear to their rotational motion, coupled with the incomplete and uncertain nature of the associated arguments suggests Matthew's construction of a mental model[212]. Additionally, the downward-gesturing of his pencil during his reference to the rider's free fall reflects his mental imagery about the rider's spatial motion under the influence of gravity. Such physical gestures mimicking spatial motion of objects have also been associated with students' mental modeling[156;213].

Matthew's Assembling of a Knowledge Framework and Noticing of an Inconsistency

On the sensemaking front, the same segment reflects the first two stages (Stage 0 and Stage 1) of the Sensemaking Epistemic Game. Matthew's initial moves – highlighting the relevant information in the problem statement and emphasizing the rider's mass – corresponds to his assembling of knowledge about the Gravitron task. Investigations on physics problem solving have noted the role of problem statements in activating students' prior knowledge[55;80;117].

Matthew's argumentative shift in the remaining part of the transcript highlights his entry condition into the Sensemaking Epistemic Game by noticing a gap in his knowledge. The perceived gap is between his existing knowledge on linear motion (ease of free fall by a heavier person) and the required knowledge on rotational motion (angular momentum) to determine the rider's status inside the Gravitron. The noticing is evidenced by the puzzling question "*is that right?*" sandwiched between the two arguments. Odden and Russ[26] note that such questions reflecting the essence of 'something not being right' mark students' entry into the Sensemaking Epistemic Game.

Summarizing the episode, Matthew's priming on the task-related information and his argumentative shift from linear to rotational motion reflect his construction of a mental model in addition to his navigation through the first two stages of the Sensemaking Epistemic Game.

6.4.2 Episode II

Matthew then argues that for a rider to be suspended inside the Gravitron, the normal force should exceed the downward pull of gravity.

> *So you need... since the rider is gonna be thrown into the wall, you need a normal force that exceeds the downward force of gravity.* [Pauses for 38 seconds.]

The prolonged pause makes the interviewer intervene with the question, "*What are you thinking?*" to which he replies as:

> *I am just trying to remember how to solve the problem.*

Constructing a force diagram, Matthew represents the forces of friction and gravity (Figure 6.5), and calculates the magnitude of the rider's centripetal acceleration by assuming that the Gravitron's walls are perpendicular to the ground.

> *I know there is... in the free body diagram, there is the downward force due to gravity* [simultaneously draws the force diagram and indicates gravity as F_g], *and then the normal force from being thrown into the wall* [draws the normal force and indicates frictional force as F_f], *and then that creates the force of friction. So, you need to solve for centripetal acceleration.*
>
> [By looking at the equation sheet] *I are given omega, so I can solve for, solve for centripetal acceleration. And then I can do F equals ma. Assuming that the wall is at ninety degree to the ground. Then... this should be correct. The centripetal*

> *acceleration is r times omega squared. So* 4.6 *times... yeah that's in radians per second. That's proper? And that's [types in calculator]* 13.8*.*

Reiterating the same assumption, he goes on to equate the denoted forces of friction and gravity. With the help of the provided equation sheet, the formulated equation is then simplified into an algebraic inequality expressed in terms of the centripetal acceleration (13.8), the coefficient of static friction (0.2), and the acceleration due to gravity (9.8) as shown in Figure 6.5.

> [Scribbling on the solution sheet] *Assuming that it's ninety degree, all of their... all of that will be converted into frictional force* [looking at the equation sheet]. *Yeah... that's the normal force.. so F equals ma. Yeah, F equals* [Pauses for 15 seconds, looks at the equation sheet]. *Is there a friction section here? Yeah* [finds the required section]. *So its the normal force into the wall, the force of friction has to equal the force of gravity, and both have mass, and this mass is negligible, so you have just... a equals... and then this a has to be bigger than or equal to* [erases].
>
> *So, the force of friction is normal force times mu equals force of gravity. Normal force in this equation is the angular acceleration times mass. For F equals ma times mu, equals mass times gravitational acceleration, and then both have mass* [cancels 'm' on both sides of the equation]. *So yeah* 13.8 *times the* 0.2*, coefficient of static friction, and that has to be greater than or equal to 9.8 for the gravitron to be able to hold people up.* [Uses calculator]

I are cautious to interpret most of the first thirteen lines (the first segment) of the above excerpt ("*Assuming that it's ninety degree [...] bigger than or equal to.*") since Matthew utters these arguments while scribbling on the solution sheet which he later erased (as noted in the transcript). Because of the lack of evidence in corroborating the student's verbal arguments with his written solution, I cannot ascertain the intended meaning of these these lines.

Table 6.3: *Summary of the features of the epistemic games in Matthew's and Ken's approaches at the Gravitron task. Both the approaches share the same target epistemic form of making a claim about the rider's status inside the Gravitron. The target epistemic form is further constrained by only two possibilities: either or not the riders would fall of the Gravitron's walls.*

Features	Matthew's approach	Ken's approach
Entry condition	Noticing of a gap between his existing knowledge on linear motion and expected knowledge on rotational motion.	Noticing of a gap in his understanding on the interplay of the forces in holding up the Gravitron's rider.
Moves	Highlights the information from the problem statement.	Identifies the information from the problem statement.
	Argues about linear and rotational motion of the Gravitron's rider.	Constructs force diagram.
	Constructs a force diagram and 'mathematizes' the diagram-based arguments.	Reasons about the association between the indicated forces in the constructed diagram.
	Interprets the physical implication of the formulated mathematical inequality to make a claim about the Gravitron's rider.	Quits the problem-solving exercise
Exit condition	Concludes that the riders would slide off the Gravitron's walls under the given conditions.	Quits the problem without making any claim.

Matthew's engagement with the Denotative Function and the Demonstration components of modeling

The current episode's transcript reflects the modeling's Denotative Function and the Demonstration components in his approach. The student focuses on the rider-Gravitron system (target system), and represents it through a force diagram highlighting the forces of friction (F_f) and gravity (F_g). Matthew's moves to represent the target system through a force diagram marks the modeling's Denotative Function in his approach. Furthermore, Matthew's moves to equate the denoted forces (i.e., mathematically relating the denoted elements of the representation), and simplifying the formulated equation into the algebraic inequality reflect the modeling's Demonstration component in his reasoning.

Matthew's Generation of an Explanation

Through the lens of sensemaking, the same segment marks Matthew's response to the perceived need to account for the rider's rotational motion as observed in the previous episode (Section 6.4.1). Matthew generates an explanation by constructing the force diagram, calculating the centripetal acceleration, and determining the inequality (in terms of the centripetal acceleration). These moves represent Matthew's navigation through the explanation generation stage (Stage 2) of the Sensemaking Epistemic Game.

In summary, Matthew's moves in this episode correspond to his engagement with the Denotative Function and the Demonstration components of modeling. The same moves also reflect his navigation through the explanation generation phase of the sensemaking process.

6.4.3 Episode III

In the end, Matthew interprets the result of the obtained algebraic inequality and concludes that riders would fall off the Gravitron's wall under the given conditions. He further goes on to claim that the Gravitron needs to be spinning at a rate of 49 radians per second for people to hold up.

> *Yeah, what I found is everyone is gonna slide because it's not spinning fast enough. And it need to be spinning* [uses calculator] *49 radians per second for people to hold up on the wall.*

As a clarifying note, I do not have evidence (either in Matthew's verbal arguments or in his written solution) reflecting the reasoning which guides the calculation of the required angular velocity of the Gravitron (49 radians per second). Consequently, the following analysis will not entail any interpretations about the same, except for acknowledging the statement.

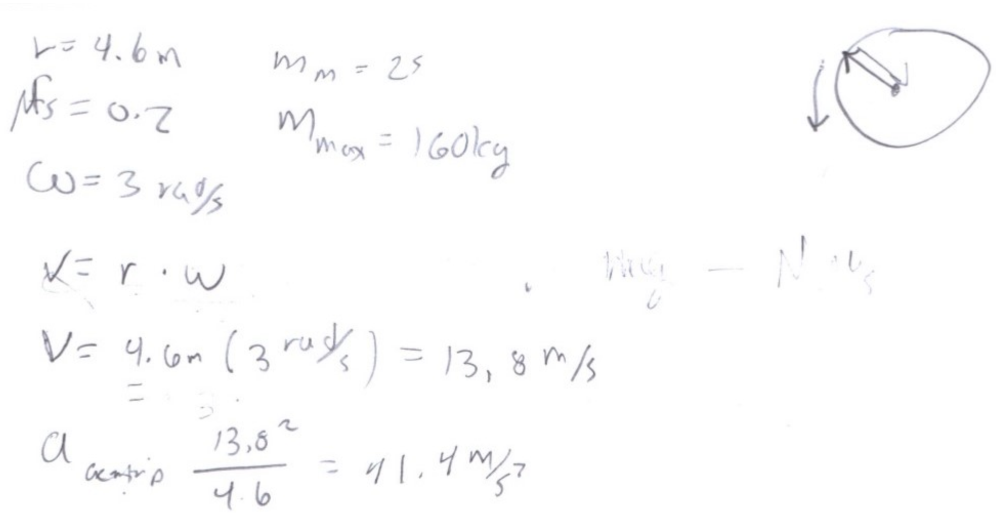
Matthew's engagement with the Inferential Function component of modeling

Matthew's final move of interpreting the physical implication of algebraic inequality through the lens of the rider-Gravitron system highlights the Inferential Function feature of modeling in his approach. The student transitions from the 'model-world' back into the 'physical world' by mapping the theoretical result (inequality) obtained in the Demonstration stage onto the target system. This mapping assists Matthew in making the required prediction thereby concluding the modeling process.

Matthew's Resolution of the perceived gap

Matthew's concluding statement also marks his exit from the sensemaking process and thus the Resolution stage of the Sensemaking Epistemic Game. He generates a coherent (mathematical) explanation by resolving his perceived need to account for the rider's rotational motion in making the required claim. He concludes clearly that riders would slide off the Gravitron's wall under the specified conditions, and further goes on to state that riders would need to be spinning at 49 radians per second to stay put. This decisive claim marks Matthew's exit condition from the Sensemaking Epistemic Game. Table 6.3 ('Matthew's approach' column) further summarizes features of epistemic games in Matthew's approach at the Gravitron task.

In summary, the concluding phase of Matthew's attempt at the Gravitron task simultaneously reflects the Inferential Function component of modeling and the Resolution stage of



$r = 4.6\,m$

$\mu_s = 0.2$

$\omega = 3\,rad/s$

$m_m = 25$

$m_{max} = 160kg$

$v = r \cdot \omega$

$mg - N\mu_s$

$V = 4.6m\,(3\,rad/s) = 13{,}8\,m/s$

$a_{centrip}\ \frac{13{,}8^2}{4.6} = 41.4\,m/s^2$

Figure 6.6: *Ken's attempt to the Gravitron Problem*

the Sensemaking Epistemic Game. Figure 6.4 (a) summarizes the relative overlap between the elements of modeling and sensemaking processes in Matthew's solution approach.

In the following section, I present another attempt at the same Gravitron task from a student named Ken (pseudonym). Unlike Matthew, Ken quits the problem-solving exercise upon making several unsuccessful bids. Similar to the analysis demonstrated in this section, I simultaneously examine Ken's problem-solving moves through the lens of modeling and sensemaking in addition to noting their relative overlap.

6.5 Ken's attempt at the Gravitron task

Upon pooling the contextual information, Ken represents the rider-Gravitron system (as shown in Figure 6.6). While reasoning about the constructed representation, Ken suggests to compare the involved forces in order to make the required claim. Unable to relate the forces, Ken discontinues his attempt, and proceeds to solve other problems from the set. Revisiting the Gravitron task, he calculates the riders' centripetal acceleration and finally abandons his pursuit. In the rest of this section, I detail Ken's attempt along with analyzing his problem-solving moves through the lens of modeling and sensemaking. Figure 6.4 (b) summarizes Ken's moves categorized into the components of modeling and stages of the

Sensemaking Epistemic Game.

6.5.1 Episode I

Ken begins by noting down the 'given' quantities from the problem statement and refers to the riders' free fall through a hand gesture.

> *So, just like off the side, its kinda like write down the stuff it gives for later use.* [Reads the problem statement for 12 seconds, and writes down the magnitudes of radius, coefficient of static friction, minimum angular velocity, minimum and maximum mass].
>
> *Okay, so ah* [looks at the equation sheet and reads the problem statement]. *I guess like in order to understand for them not to go down* [gestures his index finger downwards], *...umm..., the force... uh.*

Ken's construction of a qualitative mental model

The above transcript reflects Ken's construction of a qualitative mental model. Ken tries to articulate a condition through a hand gesture that prevents the Gravitron's riders from slipping down. He however does not complete the argument which is evidenced by the phrases "*umm... the force... uh*". The incompleteness of Ken's arguments along with the representational gesturing of the riders' spatial motion reflects his engagement with a qualitative mental model[156;213].

Ken's Assembling of Knowledge Framework

From the sensemaking perspective, the same segment marks Ken's assembling of his knowledge framework (Stage 0 of the Sensemaking Epistemic Game) evidenced by pooling of the relevant contextual information.

6.5.2 Episode II

Ken proceeds with construction of a representation as shown in Figure 6.6. He then suggests to compare the involved forces after identifying one of the indicated arrows as the force preventing the riders from slipping down. He then refers to angular acceleration amid intermittent pauses.

> *Okay so, I will draw the picture here* [draws a circle, two oppositely directed arrows and a vertically downward arrow. Reads the problem statement for 8 seconds] *And then this is...* [pointing at the outward arrow] *Uh... so this force has to keep them from falling down and then so, I need to find ah... compare, like the forces?* [Points at the diagram in the problem statement, and refers to the equation sheet for 17 seconds.]
>
> *Okay, so, my equation sheet is uhh, that alpha is delta omega by delta t and then umm...* [Pauses for 12 seconds, and looks at the problem statement and the equation sheet].

The recurring pauses makes the interviewer intervene with the question "*What are you thinking?*" to which Ken replies that he is trying to relate the information from the equation sheet to the given context.

> *Uhh, I am trying to see how I can relate uh, like what information I got to, just to see like what other information I think I kinda like to obtain to hopefully find like uh forces. I don't know if that makes any sense. Uh* [Looking towards equation sheet and pauses for 20 seconds].
>
> *I think.. Yeah, what I am trying to think is how I am gonna relate this* [pointing at the Gravitron's image from the problem statement] *to forces* [pointing at the equation sheet]. [Refers the equation sheet for 15 seconds] *Okay, so I did not find anything there.*

Ken's engagement with the Denotative Function component of modeling

This episode reflects the modeling's Denotative Function in Ken's reasoning. Ken constructs a representation by associating two radially opposite arrows and a downward arrow with forces experienced by the Gravitron's rider ("*So this force has to*[...]"). Through this representation, Ken explicitly highlights the rider-Gravitron system as the target system, and thus transitions from the physical-world to the model-world. In addition, his reply to the interviewer's question on trying to relate the information from the equation sheet to '*find forces*' suggests his attempt to associate the indicated arrows with the relevant forces. This feature of highlighting the target system through a representation, and associating features of the system (forces on the rider) through elements of the representation (arrows) marks the modeling's Denotative Function in Ken's approach.

Ken's Noticing of an Inconsistency

On the sensemaking front, the same segment reflects Ken's entry into the Sensemaking Epistemic Game through noticing of an inconsistency in his understanding of how the forces sustain the Gravitron's rider. The noticing is evidenced by (i) his suggestion to compare the forces, (ii) lack of reference to specific forces either in his verbal arguments or in his written solution, and (iii) an explicit mention of trying to relate the forces from the equation sheet ("*Yeah, what I am trying to think [...]*").

Summarizing the episode, Ken's construction of the representation, and his attempt to relate the forces from the equation sheet to the representation correspond to his engagement with the modeling's Denotative Function. The same moves also evidence his noticing of an inconsistency in his understanding on the interplay of the forces in holding up the Gravitron's rider.

6.5.3 Episode III

Struggling to locate clues from the equation sheet, Ken proceeds by reasoning about the forces acting on the rider. Noting static friction to oppose the rider's weight along the

y-direction, he refers to the relationship between friction and the normal reaction. This relationship cues Ken on calculating centripetal acceleration. Unable to reason any further, he acknowledges of being ‘stuck’ on the task, and mentions of revisiting it upon solving the remaining problems from the set.

> *Okay so, in the y direction, I have, will start with the max weight. So you have weight of person putting it down, and then static friction, would be opposing it slipping so, and static friction is normal force times* μ_s*. And then normal force is equal to...* [looks at the equation sheet]. *Okay, so I need to find alpha* [pauses for 20 seconds while looking back and forth between equation sheet and problem statement]. *So....* [pauses for 10 seconds].
>
> *Yeah, I am pretty stuck on this right now. Just gonna, I will just come back to that. There are no points involved* [smiles and turns the page for the next problem].

Ken’s engagement with the Demonstration component of modeling

Ken’s moves in this episode exemplify the modeling’s Demonstration component. Having denoted forces as arrows in the representation, Ken identifies weight, friction, and the normal force as the forces experienced by the Gravitron’s rider. He notes the force of friction to oppose the rider’s weight, thereby tying the two forces together. He then connects the friction and the normal force by citing the mathematical relation between them. Consequently, Ken’s attempt to relate the identified forces reflects the Demonstration component of modeling in his reasoning.

Ken’s Generation of an Explanation

On the sensemaking front, Ken’s reasoning about the forces marks his navigation through the penultimate stage of the Sensemaking Epistemic Game. In response to the noticed inconsistency on the interplay of the forces on the Gravitron’s rider, Ken generates an explanation by identifying the relevant forces at play and describing their interrelationships. He intuitively argues that the rider's downward motion would be opposed by the force of friction. He then invokes his formal knowledge by expressing friction as the product of normal force and the coefficient of static friction. Thus, Ken generates an explanation by blending his intuitive and formal knowledge to relate the identified forces.

Upon attempting the remaining problems, Ken returns to the Gravitron task and calculates the magnitude of centripetal acceleration. Noting the centripetal force to keep the '*kid on the ride*', he tries identifying the counter-force that works in opposition to the centripetal force.

> *I will try this one real quick. Okay.* [Looks at the equation sheet and erases what was earlier written]. *So, r times w. Radius is* 4.6 *times* 3 [Uses calculator]. *And then* [looks at the equation sheet]. *Yeah okay. So, you can find centripetal acceleration by squaring the... and dividing it by r* [uses calculator]. *So its* 44.1 *meters per second-squared. And this acceleration has to do like with the force that keeps the kid on the ride [while pointing at the radially inward arrow in the diagram]. And then okay.* [Pauses for 42 seconds] *How do I...?* [Points at the radially outward arrow in the diagram and pauses for 10 seconds]. *So, there is a force pushing back on them.* [Pauses for 22 seconds].

Ken continues to relate the forces, this time by focusing on the radial arrows indicated in his representation. He notes the radially inward arrow as the centripetal force that holds the rider inside the Gravitron ("*[...] the force that keeps the kid on the ride.*"). He then tries to reason about the oppositely directed arrow acting in opposition to the centripetal force. Ken's attempt to contrast the indicated arrows in his representation highlights his continued engagement with the Demonstration component of modeling.

Ken's determination of the centripetal acceleration and reasoning about the two forces also reflect his sustained efforts to generate explanations on the sensemaking front.

In addition, this segment also corresponds to Ken noticing an additional inconsistency on fleshing out the 'pushing back' force acting on the rider. This is evident from the question "*How do I...?*" as such vexing questions have been noted to reflect students noticing additional inconsistencies during sensemaking[26]. Ken's noticing of additional discrepancies while generating explanations is also consistent with the trajectory of the sensemaking process as highlighted by the Sensemaking Epistemic Game framework[40].

Unable to reason conceptually on the 'push back' force, Ken resorts to his memory on trying to recall a solution approach. Unable to recall a prior solution, he finally abandons his pursuit on the Gravitron task.

> *Umm... I am trying to like, umm, I am trying to recall, like the similar problem I did. There is a force...* [pauses for 22 seconds] *Well, I think that's the best I could do so far.*

This last segment of the transcript reflects Ken's unsuccessful attempt at the Demonstration component of modeling. In pursuit of determining the push-back force, Ken seeks to recall a strategy from his memory. Unable to do so, he quits the problem solving exercise.

From the sensemaking perspective, this segment marks Ken's abrupt exit from the epistemic game. Without a satisfactory explanation on the interplay between the forces in holding up Gravitron's rider, Ken gives up on his attempt after not being able to recall a solution approach from his memory. Table 6.3 (column 3) presents the features of epistemic games in Ken's attempt at the Gravitron task.

Summarizing the episode, Ken identifies weight, frictional and normal forces and relates them in the Gravitron context through his intuitive, and curricular knowledge. Struggling to make inroads, he temporarily abandons the task and revisits it. He then seeks to determine the 'push-back' force which acts in opposition to the centripetal force. Unable to make a breakthrough, the student abandons his solution pursuit. From the modeling perspective, Ken's moves in this episode reflect his struggle to relate the denoted forces in his representation, i.e., the struggle to engage in the modeling's Demonstration component. On the sensemaking front, the same moves reflect his struggle to generate a coherent explanation for his perceived gap about the forces acting on the Gravitron's rider.

6.6 Discussion

In order to probe how modeling and sensemaking intertwine, I analyzed two case studies of physics problem-solving through the lens of mental modeling, the DFDIF components, and the stages of the Sensemaking Epistemic Game (Figure 6.4). I observe that modeling – construction of mental models and engagement with DFDIF components – entails navigation through the stages of the Sensemaking Epistemic Game. Even though the components of modeling overlap with the epistemic game's stages, I do not find a one-to-one relation between the elements of the two processes. I also observe that barriers experienced in engaging with the modeling's DFDIF components can inhibit navigation through the stages of the Sensemaking Epistemic Game. These arguments are detailed below.

6.6.1 Modeling involves navigation through the stages of the Sensemaking Epistemic Game

Figure 6.4, particularly (a), indicates that the construction of a mental model and engagement with the DFDIF components involve navigation through the stages of the Sensemaking Epistemic Game. Additionally, the figure also highlights the co-occurrence of the assembling of knowledge framework stage of the epistemic game (Stage 0) with the construction of a mental model; the explanation-generation stage (Stage 2) with the Demonstration; and the Resolution stage (Stage 3) with the modeling's Inferential Function.

In the initial phases of problem-solving, both participants primed on the task-related information as part of constructing their mental models. While Matthew annotated the provided information, Ken wrote out the same explicitly. The assembling of the task-related information in both cases contributed to the participants' interaction with their mental models. My observation on the overlap between mental modeling and assembling of contextual information is consistent with the current literature, which has noted mental modeling as a key feature of making sense of the task[36] using the essential pieces of information from the problem statement[214;215].

I also note the alignment of the explanation generation during sensemaking (Stage 2) with the modeling's Demonstration feature. Both Matthew and Ken generated explanations in response to their perceived gap by mathematically relating the denoted forces in their respective representations. The alignment of the explanation generation stage with the Demonstration component comes as no surprise as numerous studies have noted sensemaking to involve coordination between multiple representations[216;217]. *Demonstration*, particularly while making sense of physics problems, can involve making effective transitions between various forms of representations such as Free Body Diagrams, graphs, equations, etc.[169;170].

Lastly, in case of Matthew, the Resolution stage of the epistemic game (since Ken did not reach this stage) aligned with the modeling's Inferential Function. Matthew concluded his sensemaking by interpreting the result of the mathematical inequality, i.e., by transitioning from the model-world to the physical-world. My observation on the overlap between the results' interpretation (Inferential Function) as the concluding part of the problem solving process (Resolution) is also congruent with the existing findings on students' use of mathematics in physics[36;55;218].

Even though the components of modeling co-occur with stages of the Sensemaking Epistemic Game, I do not observe a one-to-one correspondence between all the elements of the two processes. This primarily stems from the dynamic nature of noticing inconsistency during sensemaking. While Matthew noticed a discrepancy in his knowledge during mental modeling (Section 6.4.1 and Figure 6.4 (a)), for Ken it was associated with the Denotative Function and while *Demonstrating* the internal consistency in his representation (Section 6.5.2 & 6.5.3 and Figure 6.4(b)).

I argue that the association of modeling as a process of making sense of a phenomenon[27;35;36;38;39;46;73] can be restated more precisely as: *modeling as a process of navigating through the stages of the Sensemaking Epistemic Game when reasoning about a phenomenon.* The contemporary literature has also noted several sensemaking features, such as priming of prior knowledge, and generating explanations, as part of the modeling process[39]. My findings complement this observation by noting that these features co-occur vis-à-vis engagement with mental models and with the Demonstration feature of modeling.

6.6.2 Struggling to engage with one or more components of modeling can inhibit sensemaking

The analysis of Ken's case further reveals that barriers experienced in modeling can impede sensemaking of the given situation. From Figure 6.4 (b), I observe that the remaining facets of sensemaking (the Resolution stage) and the modeling processes (the Inferential Function component) are not evidenced as Ken chooses to abandon his attempt at the task. Despite representing the rider-Gravitron system, and reasoning about the forces acting on the Gravitron's rider, Ken's choice to give up on his approach stemmed primarily from his repeated unsuccessful attempts to relate the two oppositely directed arrows in his representation. Upon noting the centripetal force (the radially inward arrow in his representation), Ken struggled to reason about the 'push-back' force (the radially outward arrow) which acted in opposition to the identified centripetal force. From the modeling perspective, Ken's struggle tends to be associated with relating the elements (forces) of the constructed representation, i.e., the Demonstration component of modeling. The hindrance in the modeling's Demonstration component resulted in the lack of a satisfactory explanation to his perceived gap, which in turn nudged him to quit the sensemaking process.

I emphasize that Ken's decision to give up on his attempt may not have been caused *solely* by the modeling barriers. The lack of conceptual coherence in his arguments, or his 'working-model' of the Gravitron may also have played a crucial role in nudging him to quit sensemaking. However, I prioritize modeling barriers over other factors since modeling or sensemaking can occur independently of conceptual coherence. Other participants in the data made the requisite claim upon making sense of or modeling the Gravitron scenario, albeit their conceptual correctness can be contested. Furthermore, theobservation on modeling barriers impeding sensemaking addresses the contemporary literature's call to investigate the contextual factors which encourage students to give up on sensemaking[40].

Chapter 7

Conclusion and Future Work

Contributing to the contemporary efforts of the education research community, this dissertation discusses ways of leveraging existing theoretical ideas in developing and analyzing assessments that promote sensemaking in physics. Arguments discussed in this work can assist educators in taking a step closer towards promoting a set of valued objectives in science learning environments such as sophisticated epistemology, enhanced content understanding, and students' ability to generate new knowledge.

Chapter 3 presented a broad overview of the development process of a novel research-based assessment that facilitates students' sensemaking of their surrounding world using disciplinary ideas. Specifically, I elucidated the leveraging of existing theoretical and design frameworks in the development of the Thermal and Statistical Physics Assessment (TaSPA). This work presents a paradigm shift in how assessments are envisioned and designed by the discipline-based education research community. I presented the TaSPA's "theory-of-action"- an explicit account of how the assessment can bring about an intended change by shifting the focus of our classrooms from "knowing" to "doing" science.

Furthermore, the contemporary literature has raised several concerns on the existing assessments which include (i) lack of clarity for instructors in interpreting students' assessment scores, (ii) the need to shift the focus of assessments from what students know to what students can do with what they know ('knowledge-in-use'), (iii) misalignment between the assessment goals and the valued objectives of the instructors, and lastly (iv) scarcity of standardized assessments for undergraduate thermal and statistical physics courses. TaSPA addresses the first concern by providing actionable feedback for instructors on how they can go about modifying their courses based on their students' responses to the assessment tasks. Additionally, by blending scientific practices with the conceptual ideas of thermal and statistical physics, TaSPA focuses on assessing the 'doing' aspect of science, thereby addressing the second concern. Since instructors get to choose what they wish to assess (by choosing the learning performances they value), the current assessment contributes in minimizing the existing gap between the objectives of assessment and local goals of instructors (thereby addressing the remaining two concerns).

As for the future work, theoretical underpinnings in translating and interpreting the individual responses to the class' results are being explored. The research team is also in the process of piloting the entire exercise of assessment administration, starting from instructors choosing learning performances to getting actionable feedback based on their students' responses at numerous undergraduate institutions in the United States. The assessment is being planned to be available to the public in 2024.

In Chapter 4, I put forward a theoretical argument that to promote sensemaking, tasks should cue students to unpack the underlying mechanism of a real-world phenomenon by coordinating multiple representations and by physically interpreting mathematical expressions. I make this argument by leveraging existing theoretical perspectives on the cognitive features of sensemaking, and by adopting conjecture mapping[1].

One of the primary contributions of this work involves adopting an agent-based approach in articulating task-design arguments in physics. Research on task design has traditionally focused on the valued objectives of researchers by overlooking the role of contextual features in influencing students' engagement with tasks. Chapter 4 presents an exemplar case by simultaneously attending to the valued objectives of the researchers along with accounting for the students contextual factors. While my theoretical approach on deducing sensemaking elements from its definition reflects the researchers' valued objectives in task design, the vocabulary adopted in making task-related arguments reflect the consideration of students' contextual factors.

This work contributes to the contemporary literature by operationalizing conjecture mapping in the context of task design in physics. This technique has been traditionally employed in designing learning environments such as (but not limited to) vocational training[219], online or hybrid learning[220;221], or pedagogy in informal communities[222]. Chapter 4 leverages this framework in designing physics tasks. Operationalization of this framework also brings together the broad literature on sensemaking. While the "embodiment" and "mediating processes" (Section 4.1.2) encompass the theoretical and analytical views on sensemaking, the "outcomes" embodies the literature on the effects of sensemaking.

For instructors, this chapter provides a generalized framework for designing assessments, or crafting examples for classroom discussions that can promote sensemaking. The generalized nature of these task features provide avenues for instructors to design tasks based on their local learning objectives and curricula, thereby facilitating their agency[115;223].

For researchers, the current work describes a methodology for identifying task features, which can be extended to promote other valued epistemic practices such as argumentation or modeling[46]. My proposed methodology – extracting salient features of a cognitive process from its definition, and back-tracking the task characteristics – can contribute to the community's efforts in promoting valued epistemic practices in our classrooms. Additionally, there is an increasing traction of investigations on sensemaking in laboratories[224;225]. Researchers can also extend the presented methodology in identifying features of activities or tasks that promote sensemaking during experimentation.

Since I do not claim the identified features to be exhaustive, researchers can expand on the proposed list, or can further investigate the conditions in which the identified task features are effective. Contemporary reports on the proposed task features do indicate certain accompanying constraints. Heckler[121] notes explicit prompting on constructing representations may cue protocol-based approaches to learning as opposed to intuitive engagement with the content. Similarly, researchers have noted real-world contexts in tasks to elicit subjective judgements about the scenarios[188], and initiating gender-based disparity in performances[226].

This study also opens up avenues to explore the interaction of the proposed task features with their structural features. Research on task design has noted activities such as designing experiments, or modeling complex systems to differ from solving the typical end-of-the chapter physics problems. Unlike the former (referred as ill-structured problems), the latter (well-structured) tasks have a well-defined protocol for initiating, proceeding, and terminating the activity[227–231]. The current study paves way for researchers to probe the influence of our identified task features in the context of well- and ill-structured problems.

In Chapter 5, I focused on the "analysis" component of my dissertation by going through students' responses to tasks embodying features articulated in Chapter 4. I present a qualitative proof-of-concept of the likelihood argument about the task features by analyzing reasoning approaches from two participants (Luke and Catherine) about a same task (Figure 5.2). While Luke makes the required claim by engaging in sensemaking, Catherine makes the claim through "plug-and-chug" approach. Furthermore, through Luke's case study, I present detailed evidence about how a specific proposed task-feature can elicit a specific element of sensemaking.

In Chapter 6, I analyze two additional problem-solving approaches on another task embodying the features proposed in Chapter 4. This work explicitly demonstrates the intertwining between sensemaking and modeling by employing the Sensemaking Epistemic Game[40], mental modeling and the DFDIF account of modeling[4]. Qualitative analysis of the participants' problem-solving reveals that modeling – constructing a mental model and engaging with the DFDIF components – entails navigation through the four stages of the Sensemaking Epistemic Game. In addition, the analysis of Ken's approach reveals that barriers experienced in modeling can influence students to prematurely quit sensemaking.

Observations made in this chapter provide insights for instructors for making potential interventions when students tend to give up on sensemaking. For instance, a useful intervention in the case of Ken, who abandoned sensemaking upon struggling with modeling's Demonstration component, might have been to prompt him to reflect on how his identified forces – gravity, friction and normal force – connect to the denoted centripetal force. I suspect students can also experience barriers with other modeling features such as construction of mental models, and engagement with the Denotative Function, and the Inferential

Function components (though not observed in our case studies). Students' initial mental models particularly while reasoning about a real-world context like the Gravitron, are associated with ideas about the functioning of their surrounding world. Potential interventions on mental modeling might include prompting students to explicitly articulate their intuitive ideas and discussing about how those ideas manifest as scientific principles in the given context. As for modeling's Denotative Function, the interventions can include explicitly making the "epistemic status" of denoting physical quantities through algebraic symbols as an arbitrary choice often devoid of physical or mathematical implications (e.g., denoting Friction as "F_f" or "f" has no bearing on physics or mathematical arguments being made in a given context)[17;19]. Concerning the Inferential Function, explicit emphasis about reflecting on the "physical meaning" of the obtained mathematical results, or modifying known equations in light of given physical conditions can assist students in sensemaking during problem solving.

For researchers, the contextual operationalization of the Sensemaking Epistemic Game contributes a new theoretical perspective to analyze students' sensemaking on physics problems. Even though researchers have probed students' reasoning using the epistemic games construct, the focus has primarily been on the use of mathematics in physics[82] or "answer-making" during problem solving[232]. Students' sensemaking using mathematical formalisms or 'mathematical sensemaking' has been investigated through various constructs such as mediated cognition[75], blended processing[90], symbolic forms[92], etc. The current work adds the Sensemaking Epistemic Game to this list of frameworks employed to probe students' sensemaking during problem solving. Lastly, consistent with the current arguments in the modeling literature[35;206], this work takes the agent-based perspective to analyze students' modeling. That is, I investigate modeling through the cognitive interests and aims of the participant (models-for) rather than considering their mere representation of the concerned context (models-of)[206].

7.0.1 Limitations

A number of caveats accompany the claims made in this dissertation. Firstly, the sensemaking process is driven by a perceived gap in one's knowledge system[21]. This noticing of inconsistencies depends on prior knowledge, awareness, self-evaluation, and approaches students employ while reasoning about the given scenario[125]. The proposed task features in Chapter 4 do not attend to this crucial element of the sensemaking process due to its contextual and meta-cognitive nature. Additionally, my objective is neither to argue that the proposed task features *necessarily* engage students into sensemaking, nor these features to be the *only ones* to promote sensemaking. Rather, I make a modest argument that the identified features which when present in a task *together*, enhance the likelihood of students sensemaking in physics.

Chapter 6 accompanies several limitations. The DFDIF account, which has been used as the modeling framework, captures the "big-picture" of the modeling process by categorizing it into three broad activities: the Denotative Function, Demonstration, and Inferential Function. Modeling is a complex iterative process entailing activities such as making assumptions, approximations, etc. and the adopted framework does not capture these aspects. The arguments made in this study are based on the analysis of responses of two white, male students on a single problem. I acknowledge that a person's view of the outside world is informed by their positionality, and the claims made in this paper do not take this factor into account. Analysis of demographically diverse students' approaches to a wide range of problems would undoubtedly enrich the claims. And lastly, while my data includes participants thinking aloud, there are moments especially during prolonged pauses which involve interventions from the interviewer asking participants to articulate their thoughts. Though an inherent limitation of the think-aloud protocol, the interviewer's interjection does have an impact on the participants' thought process.

As part of future work, I seek to extend our observations to identify the importance of the constructed inequality in Matthew's solution to both his modeling and sensemaking. Additionally, I also seek to investigate whether the converse of our second claim holds true, i.e., whether barriers experienced in sensemaking can inhibit modeling of the given context. Such explorations can further the pedagogical efforts in supporting students' modeling and sensemaking thereby making classroom experiences more exploratory.

Bibliography

[1] William Sandoval. Conjecture mapping: An approach to systematic educational design research. *Journal of the learning sciences*, 23(1):18–36, 2014.

[2] Galileo Galilei. Two new sciences (s. drake, trans.). *Madison: University of Wisconsin Press.(Original work published 1638)*, 1974.

[3] Richard IG Hughes. Models and representation. *Philosophy of science*, 64:S325–S336, 1997.

[4] Mauricio Suárez. Scientific representation, denotation, and fictional entities. In *Recent developments in the philosophy of science: EPSA13 Helsinki*, pages 331–341. Springer, 2015.

[5] James G Greeno. The situativity of knowing, learning, and research. *American psychologist*, 53(1):5, 1998.

[6] Jean Lave. *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge University Press, 1988.

[7] Adrian Madsen, Sarah B McKagan, and Eleanor C Sayre. Resource letter rbai-1: research-based assessment instruments in physics and astronomy. *American Journal of Physics*, 85(4):245–264, 2017.

[8] Jennifer L Docktor and José P Mestre. Synthesis of discipline-based education research in physics. *Physical Review Special Topics-Physics Education Research*, 10(2):020119, 2014.

[9] Adrian Madsen, Sarah B McKagan, and Eleanor C Sayre. Best practices for administering concept inventories. *The Physics Teacher*, 55(9):530–536, 2017.

[10] Adrian Madsen, Sarah B McKagan, Mathew Sandy Martinuk, Alexander Bell, and Eleanor C Sayre. based assessment affordances and constraints: Perceptions of physics faculty. *Physical Review Physics Education Research*, 12(1):010115, 2016.

[11] Katherine D Rainey, Michael Vignal, and Bethany R Wilcox. Designing upper-division thermal physics assessment items informed by faculty perspectives of key content coverage. *Physical Review Physics Education Research*, 16(2):020113, 2020.

[12] Christopher J Harris, Joseph S Krajcik, James W Pellegrino, and Angela Haydel DeBarger. Designing knowledge-in-use assessments to promote deeper learning. *Educational measurement: issues and practice*, 38(2):53–67, 2019.

[13] "". *Next Generation Science Standards: For States, By States.* The National Academies Press, Washington, DC, 2013. ISBN 978-0-309-27227-8. doi: 10.17226/18290. URL https://www.nap.edu/catalog/18290/next-generation-science-standards-for-states-by-states.

[14] James Laverty, Amogh Sirnoorkar, Amali Jambuge, Katherine Rainey, Weaver Joshua, Alexander Adamson, and Bethany Wilcox. A new paradigm for research-based assessment development. In *Physics Education Research Conference 2022*, pages 279–284, 2022.

[15] Andrew Elby and David Hammer. On the substance of a sophisticated epistemology. *Science education*, 85(5):554–567, 2001.

[16] Thomas J Bing and Edward F Redish. Epistemic complexity and the journeyman-expert transition. *Physical Review Special Topics-Physics Education Research*, 8(1): 010105, 2012.

[17] Amogh Sirnoorkar, Anwesh Mazumdar, and Arvind Kumar. Students' epistemic understanding of mathematical derivations in physics. *European Journal of Physics*, 38 (1):015703, 2016.

[18] Amogh Sirnoorkar, Christopher AF Hass, Qing X Ryan, and Eleanor C Sayre. Qualitative analysis of students' epistemic framing surrounding instructor's interaction. In *Proceedings of the Physics Education Research Conference (PERC*, pages 478–483, 2020.

[19] Amogh Sirnoorkar, Anwesh Mazumdar, and Arvind Kumar. Towards a content-based epistemic measure in physics. *Physical Review Physics Education Research*, 16(1): 010103, 2020.

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

[21] Tor Ole B Odden and Rosemary S Russ. Defining sensemaking: Bringing clarity to a fragmented theoretical construct. *Science Education*, 103(1):187–205, 2019.

[22] Matthew A Cannady, Paulette Vincent-Ruz, Joo Man Chung, and Christian D Schunn. Scientific sensemaking supports science content learning across disciplines and instructional contexts. *Contemporary Educational Psychology*, 59:101802, 2019.

[23] Michael J Ford. A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, 30(3):207–245, 2012.

[24] Leema Kuhn Berland and Brian J Reiser. Making sense of argumentation and explanation. *Science education*, 93(1):26–55, 2009.

[25] Brian A Danielak, Ayush Gupta, and Andrew Elby. Marginalized identities of sensemakers: Reframing engineering student retention. *Journal of Engineering Education*, 103(1):8–44, 2014.

[26] Tor Ole B Odden and Rosemary S Russ. Vexing questions that sustain sensemaking. *International Journal of Science Education*, 41(8):1052–1070, 2019.

[27] Rosemary S Russ and Tor Ole B Odden. Intertwining evidence-and model-based

reasoning in physics sensemaking: An example from electrostatics. *Physical Review Physics Education Research*, 13(2):020105, 2017.

[28] Odd Sand, Tor Ole Odden, Christine Lindstrom, and Marcos Caballero. How computation can facilitate sensemaking about physics: A case study. In *Proceedings, Physics Education Research Conference*, 01 2019. doi: 10.1119/perc.2018.pr.Sand.

[29] 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):020109, 2020.

[30] Robert J Mislevy and Geneva D Haertel. Implications of evidence-centered design for educational testing. *Educational measurement: issues and practice*, 25(4):6–20, 2006.

[31] Deborah L Butler and Philip H Winne. Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, 65(3):245–281, 1995.

[32] Amogh Sirnoorkar and James T Laverty. Theoretical exploration of task features that facilitate student sensemaking in physics. *arXiv preprint arXiv:2302.11478*, 2023.

[33] Amogh Sirnoorkar and James Laverty. A methodology for identifying task features that facilitate sensemaking. In *Physics Education Research Proceedings*, pages 396–401, 2021.

[34] Amogh Sirnoorkar, Paul DO Bergeron, and James T Laverty. Sensemaking and scientific modeling: Intertwined processes analyzed in the context of physics problem solving. *Physical Review Physics Education Research*, 19(1):010118, 2023.

[35] Cynthia Passmore, Julia Svoboda Gouvea, and Ronald Giere. Models in science and in learning science: Focusing scientific practice on sense-making. In *International handbook of research in history, philosophy and science teaching*, pages 1171–1202. Springer, 2014.

[36] David Sands. Modeling as sensemaking: towards a theory of modelling in physics education. *European Journal of Physics*, 42(6):064001, 2021.

[37] Paul DO Bergeron, Amogh Sirnoorkar, and James T Laverty. Towards an analytic framework for characterizing student use of models. In *Proceedings of the Physics Education Research Conference (PERC*, pages 45–50, 2021.

[38] Christina V Schwarz, Brian J Reiser, Elizabeth A Davis, Lisa Kenyon, Andres Achér, David Fortus, Yael Shwartz, Barbara Hug, and Joe Krajcik. Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 46(6):632–654, 2009.

[39] Ying-Chih Chen. Epistemic uncertainty and the support of productive struggle during scientific modeling for knowledge co-development. *Journal of Research in Science Teaching*, 59(3):383–422, 2022.

[40] Tor Ole B Odden and Rosemary S Russ. Sensemaking epistemic game: A model of student sensemaking processes in introductory physics. *Physical Review Physics Education Research*, 14(2):020122, 2018.

[41] Timothy J Nokes-Malach and Jose P Mestre. Toward a model of transfer as sensemaking. *Educational Psychologist*, 48(3):184–207, 2013.

[42] Ileana María Greca and Marco Antonio Moreira. Mental, physical, and mathematical models in the teaching and learning of physics. *Science Education*, 86(1):106–121, 2002. ISSN 1098-237X. doi: 10.1002/sce.10013. URL http://onlinelibrary.wiley.com/doi/abs/10.1002/sce.10013.

[43] FangFang Zhao and Anita Schuchardt. Development of the sci-math sensemaking framework: categorizing sensemaking of mathematical equations in science. *International Journal of STEM Education*, 8(1):1–18, 2021.

[44] Lorenzo Magnani and Tommaso Bertolotti. *Springer handbook of model-based science.* Springer, 2017.

[45] Dedre Gentner and Albert L Stevens. *Mental models.* Psychology Press, 2014.

[46] National Research Council et al. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* National Academies Press, 2012.

[47] John R Frederiksen, Barbara Y White, and Joshua Gutwill. Dynamic mental models in learning science: The importance of constructing derivational linkages among models. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 36(7):806–836, 1999.

[48] Lei Bao and Edward F Redish. Model analysis: Representing and assessing the dynamics of student learning. *Physical Review Special Topics-Physics Education Research*, 2 (1):010103, 2006.

[49] Claudio Fazio, Onofrio Rosario Battaglia, and Benedetto Di Paola. Investigating the quality of mental models deployed by undergraduate engineering students in creating explanations: The case of thermally activated phenomena. *Physical Review Special Topics-Physics Education Research*, 9(2):020101, 2013.

[50] Nilüfer Didiş Körhasan, Ali Eryılmaz, and Şakir Erkoç. The influence of instructional interactions on students' mental models about the quantization of physical observables: a modern physics course case. *European Journal of Physics*, 37(1):015701, 2015.

[51] Hans Bekaert, Hans Van Winckel, Wim Van Dooren, An Steegen, and Mieke De Cock. Identifying students' mental models of the apparent motion of the sun and stars. *Physical Review Physics Education Research*, 18(1):010130, 2022.

[52] David Hestenes. Toward a modeling theory of physics instruction. *American journal of physics*, 55(5):440–454, 1987.

[53] Ibrahim Halloun. Schematic modeling for meaningful learning of physics. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 33(9):1019–1041, 1996.

[54] Andrew Pawl, Analia Barrantes, and David E Pritchard. Modeling applied to problem solving. In *AIP Conference Proceedings*, volume 1179, pages 51–54. American Institute of Physics, 2009.

[55] 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):020119, 2013.

[56] Edward F Redish and Karl A Smith. Looking beyond content: Skill development for engineers. *Journal of Engineering Education*, 97(3):295–307, 2008.

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

[58] Eugenia Etkina, Aaron Warren, and Michael Gentile. The role of models in physics instruction. *The physics teacher*, 44(1):34–39, 2006.

[59] Malcolm Wells, David Hestenes, and Gregg Swackhamer. A modeling method for high school physics instruction. *American journal of physics*, 63(7):606–619, 1995.

[60] Vashti Sawtelle, Eric Brewe, and Laird H Kramer. Positive impacts of modeling instruction on self-efficacy. In *AIP Conference Proceedings*, volume 1289, pages 289–292. American Institute of Physics, 2010.

[61] Eric Brewe, Vashti Sawtelle, Laird H Kramer, George E O'Brien, Idaykis Rodriguez, and Priscilla Pamelá. Toward equity through participation in modeling instruction in introductory university physics. *Physical Review Special Topics-Physics Education Research*, 6(1):010106, 2010.

[62] Ibrahim Abou Halloun and David Hestenes. Modeling instruction in mechanics. *American Journal of Physics*, 55(5):455–462, 1987.

[63] Eric Brewe, Laird Kramer, and George O'Brien. Modeling instruction: Positive attitudinal shifts in introductory physics measured with class. *Physical Review Special Topics-Physics Education Research*, 5(1):013102, 2009.

[64] Eric Brewe, Adrienne Traxler, Jorge De La Garza, and Laird H Kramer. Extending positive class results across multiple instructors and multiple classes of modeling instruction. *Physical Review Special Topics-Physics Education Research*, 9(2):020116, 2013.

[65] Daryl McPadden and Eric Brewe. Impact of the second semester university modeling instruction course on students' representation choices. *Physical Review Physics Education Research*, 13(2):020129, 2017.

[66] Dimitri R Dounas-Frazer and HJ Lewandowski. The modelling framework for experimental physics: Description, development, and applications. *European Journal of Physics*, 39(6):064005, 2018.

[67] Ismo Tapio Koponen. Models and modelling in physics education: A critical re-analysis of philosophical underpinnings and suggestions for revisions. *Science & Education*, 16 (7):751–773, 2007.

[68] Benjamin M Zwickl, Dehui Hu, Noah Finkelstein, and HJ Lewandowski. Model-based reasoning in the physics laboratory: Framework and initial results. *Physical Review Special Topics-Physics Education Research*, 11(2):020113, 2015.

[69] Dimitri R Dounas-Frazer, Kevin L Van De Bogart, MacKenzie R Stetzer, and HJ Lewandowski. Investigating the role of model-based reasoning while troubleshooting an electric circuit. *Physical Review Physics Education Research*, 12(1):010137, 2016.

[70] Laura Ríos, Benjamin Pollard, Dimitri R Dounas-Frazer, and HJ Lewandowski. Using think-aloud interviews to characterize model-based reasoning in electronics for a laboratory course assessment. *Physical Review Physics Education Research*, 15(1):010140, 2019.

[71] Jacob T Stanley, Weifeng Su, and HJ Lewandowski. Using lab notebooks to examine students' engagement in modeling in an upper-division electronics lab course. *Physical Review Physics Education Research*, 13(2):020127, 2017.

[72] Benjamin M Zwickl, Noah Finkelstein, and HJ Lewandowski. Incorporating learning goals about modeling into an upper-division physics laboratory experiment. *American Journal of Physics*, 82(9):876–882, 2014.

[73] Christina V Schwarz, Cynthia Passmore, and Brian J Reiser. *Helping students make sense of the world using next generation science and engineering practices*. NSTA Press, 2017.

[74] David Sands. Modeling as sensemaking: towards a theory of modelling in physics education. *European Journal of Physics*, 2020.

[75] Julian D Gifford and Noah D Finkelstein. Categorical framework for mathematical sense making in physics. *Physical Review Physics Education Research*, 16(2):020121, 2020.

[76] Michael Ford. Disciplinary authority and accountability in scientific practice and learning. *Science Education*, 92(3):404–423, 2008.

[77] Michael Ford. 'grasp of practice'as a reasoning resource for inquiry and nature of science understanding. *Science & Education*, 17(2-3):147–177, 2008.

[78] Dina Zohrabi Alaee. *Processes of problem-solving and instructional change in physics*. PhD thesis, "", 2020.

[79] Paul Hutchison and David Hammer. Attending to student epistemological framing in a science classroom. *Science Education*, 94(3):506–524, 2010.

[80] David Hammer, Andrew Elby, Rachel E Scherr, and Edward F Redish. Resources, framing, and transfer. *Transfer of learning from a modern multidisciplinary perspective*, 89, 2005.

[81] Allan Collins and William Ferguson. Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational psychologist*, 28(1):25–42, 1993.

[82] 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):020101, 2007.

[83] Seth Rosenberg, David Hammer, and Jessica Phelan. Multiple epistemological coherences in an eighth-grade discussion of the rock cycle. *The Journal of the Learning Sciences*, 15(2):261–292, 2006.

[84] Shulamit Kapon. Unpacking sensemaking. *Science Education*, 101(1):165–198, 2017.

[85] Paul J Emigh, Jonathan W Alfson, and Elizabeth Gire. Student sensemaking about equipotential graphs. In *Proceedings of the Physics Education Research Conference (PERC*, pages 99–102, 2019.

[86] MacKenzie Lenz. Sensemaking throughout the physics curriculum: Understanding expert and student ideas about sensemaking in a physics context. *""*, 2020.

[87] Benjamin W Dreyfus, Andrew Elby, Ayush Gupta, and Erin Ronayne Sohr. Mathematical sense-making in quantum mechanics: An initial peek. *Physical Review Physics Education Research*, 13(2):020141, 2017.

[88] Kelby T Hahn, Paul J Emigh, MacKenzie Lenz, and Elizabeth Gire. Student sensemaking on homework in a sophomore mechanics course. In *PERC Proceedings*, pages 26–27, 2017.

[89] Kelby T Hahn, Paul J Emigh, and Elizabeth Gire. Sensemaking in special relativity: developing new intuitions. In *Physics Education Research Conference 2019-PER Conference series*, pages 196–201, 2019.

[90] 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, 2013.

[91] Ayush Gupta and Andrew Elby. Beyond epistemological deficits: Dynamic explanations of engineering students' difficulties with mathematical sense-making. *International Journal of Science Education*, 33(18):2463–2488, 2011.

[92] Bruce L Sherin. How students understand physics equations. *Cognition and instruction*, 19(4):479–541, 2001.

[93] Christine Chin and David E Brown. Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 37(2):109–138, 2000.

[94] Maria Ruibal-Villasenor, Eugenia Etkina, Anna Karelina, David Rosengrant, Rebecca Jordan, and Alan Van Heuvelen. From physics to biology: Helping students attain all-terrain knowledge. In *AIP Conference Proceedings*, volume 951, pages 96–99. American Institute of Physics, 2007.

[95] Shulamit Kapon and Andrea A diSessa*. Reasoning through instructional analogies. *Cognition and Instruction*, 30(3):261–310, 2012.

[96] Christina Krist, Christina V Schwarz, and Brian J Reiser. Identifying essential epistemic heuristics for guiding mechanistic reasoning in science learning. *Journal of the Learning Sciences*, 28(2):160–205, 2019.

[97] Rachel E Scherr and David Hammer. Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2):147–174, 2009.

[98] Andrew Elby and David Hammer. Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. *Personal epistemology in the classroom: Theory, research, and implications for practice*, 4(1):409–434, 2010.

[99] Andrea A DiSessa. Toward an epistemology of physics. *Cognition and instruction*, 10 (2-3):105–225, 1993.

[100] Amogh Sirnoorkar, James T Laverty, and Paul Bergeron. Sensemaking and scientific modeling: Intertwined processes analyzed in the context of physics problem solving. *arXiv preprint arXiv:2207.03939*, 2022.

[101] Chauncey Monte-Sano. Beyond reading comprehension and summary: Learning to read and write in history by focusing on evidence, perspective, and interpretation. *Curriculum Inquiry*, 41(2):212–249, 2011.

[102] Gina N Cervetti, Anne L DiPardo, and Sara J Staley. Entering the conversation: Exploratory talk in middle school science. *the elementary school journal*, 114(4):547–572, 2014.

[103] Miranda Suzanne Fitzgerald and Annemarie Sullivan Palincsar. Teaching practices that support student sensemaking across grades and disciplines: A conceptual review. *Review of Research in Education*, 43(1):227–248, 2019.

[104] Thomas J Bing and Edward F Redish. Analyzing problem solving using math in physics: Epistemological framing via warrants. *Physical Review Special Topics-Physics Education Research*, 5(2):020108, 2009.

[105] Michael Kane. Certification testing as an illustration of argument-based validation. *Measurement*, 2(3):135–170, 2004.

[106] Randy Elliot Bennett. Cognitively based assessment of, for, and as learning (cbal): A preliminary theory of action for summative and formative assessment. *Measurement*, 8(2-3):70–91, 2010.

[107] James T Laverty and Marcos D Caballero. Analysis of the most common concept inventories in physics: What are we assessing? *Physical Review Physics Education Research*, 14(1):010123, 2018.

[108] James T Laverty, Sonia M Underwood, Rebecca L Matz, Lynmarie A Posey, Justin H Carmel, Marcos D Caballero, Cori L Fata-Hartley, Diane Ebert-May, Sarah E Jardeleza, and Melanie M Cooper. Characterizing college science assessments: The three-dimensional learning assessment protocol. *PloS one*, 11(9):e0162333, 2016.

[109] D Royce Sadler. Formative assessment and the design of instructional systems. *Instructional science*, 18(2):119–144, 1989.

[110] Bethany R Wilcox and Steven J Pollock. Coupled multiple-response versus free-response conceptual assessment: An example from upper-division physics. *Physical Review Special Topics-Physics Education Research*, 10(2):020124, 2014.

[111] Katherine D Rainey, Amali Priyanka Jambuge, James T Laverty, and Bethany R Wilcox. Developing coupled, multiple-response assessment items addressing scientific practices. In *Proceedings of the Physics Education Research Conference*, 2020.

[112] Hannu Nurmi. Voting theory. *e-Democracy: A Group Decision and Negotiation Perspective*, pages 101–123, 2010.

[113] Julian D. Gifford and Noah Finkelstein. A framework for curriculum design to support mathematical sense making. In *Physics Education Research Conference 2020*, PER Conference, pages 173–178, Virtual Conference, July 22-23 2020.

[114] David Stroupe, Jean Moon, and Sarah Michaels. Introduction to special issue: Epistemic tools in science education. *Science Education*, 103(4):948–951, 2019.

[115] Emily Miller, Eve Manz, Rosemary Russ, David Stroupe, and Leema Berland. Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, 55(7):1053–1075, 2018.

[116] Margaret Vaughn. What is student agency and why is it needed now more than ever? *Theory Into Practice*, 59(2):109–118, 2020.

[117] Kelli Shar, Rosemary S Russ, and James T Laverty. Student epistemological framing on paper-based assessments. *Physical Review Physics Education Research*, 16(2):020101, 2020.

[118] Kirsten Berthold, Heidi Röder, Daniel Knörzer, Wolfgang Kessler, and Alexander Renkl. The double-edged effects of explanation prompts. *Computers in Human Behavior*, 27(1):69–75, 2011.

[119] Kirsten Berthold and Alexander Renkl. Instructional aids to support a conceptual understanding of multiple representations. *Journal of educational psychology*, 101(1): 70, 2009.

[120] Timothy J Nokes, Robert GM Hausmann, Kurt VanLehn, and Sophia Gershman. Testing the instructional fit hypothesis: the case of self-explanation prompts. *Instructional Science*, 39:645–666, 2011.

[121] Andrew F Heckler. Some consequences of prompting novice physics students to construct force diagrams. *International Journal of Science Education*, 32(14):1829–1851, 2010.

[122] Cornelia S Große and Alexander Renkl. Effects of multiple solution methods in mathematics learning. *Learning and Instruction*, 16(2):122–138, 2006.

[123] Bethany Rittle-Johnson and Abbey M Loehr. Eliciting explanations: Constraints on when self-explanation aids learning. *Psychonomic bulletin & review*, 24:1501–1510, 2017.

[124] Leema K Berland, Christina V Schwarz, Christina Krist, Lisa Kenyon, Abraham S Lo, and Brian J Reiser. Epistemologies in practice: Making scientific practices meaningful for students. *Journal of Research in Science Teaching*, 53(7):1082–1112, 2016.

[125] Ross C Anderson, Shawn Irvin, Tracy Bousselot, Nate Beard, and Paul Beach. Grasping the uncertainty of scientific phenomena: A creative, agentic, and multimodal model for sensemaking. In *Uncertainty: A Catalyst for Creativity, Learning and Development*, pages 159–179. Springer, 2022.

[126] Patrick Löffler, Marcela Pozas, and Alexander Kauertz. How do students coordinate context-based information and elements of their own knowledge? an analysis of students' context-based problem-solving in thermodynamics. *International Journal of Science Education*, 40(16):1935–1956, 2018.

[127] Eric Eich. Context, memory, and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4):764, 1985.

[128] Steven M Smith, Arthur Glenberg, and Robert A Bjork. Environmental context and human memory. *Memory & Cognition*, 6(4):342–353, 1978.

[129] Duncan R Godden and Alan D Baddeley. Context-dependent memory in two natural environments: On land and underwater. *British Journal of psychology*, 66(3):325–331, 1975.

[130] Keith A Hutchison, David A Balota, James H Neely, Michael J Cortese, Emily R Cohen-Shikora, Chi-Shing Tse, Melvin J Yap, Jesse J Bengson, Dale Niemeyer, and Erin Buchanan. The semantic priming project. *Behavior research methods*, 45(4): 1099–1114, 2013.

[131] Ludovic Ferrand and Boris New. Semantic and associative. *Mental lexicon: Some words to talk about words*, page 25, 2004.

[132] Allan M Collins and Elizabeth F Loftus. A spreading-activation theory of semantic processing. *Psychological review*, 82(6):407, 1975.

[133] Margareta Enghag, Peter Gustafsson, and Gunnar Jonsson. From everyday life experiences to physics understanding occurring in small group work with context rich

problems during introductory physics work at university. *Research in Science Education*, 37(4):449–467, 2007.

[134] Kevin P Madore, Donna Rose Addis, and Daniel L Schacter. Creativity and memory: Effects of an episodic-specificity induction on divergent thinking. *Psychological science*, 26(9):1461–1468, 2015.

[135] Kevin P Madore, Helen G Jing, and Daniel L Schacter. Divergent creative thinking in young and older adults: Extending the effects of an episodic specificity induction. *Memory & cognition*, 44(6):974–988, 2016.

[136] Benjamin C Storm and Trisha N Patel. Forgetting as a consequence and enabler of creative thinking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6):1594, 2014.

[137] Bernard A Nijstad, Wolfgang Stroebe, and Hein FM Lodewijkx. Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of experimental social psychology*, 38(6):535–544, 2002.

[138] Tim George and Jennifer Wiley. Need something different? here's what's been done: Effects of examples and task instructions on creative idea generation. *Memory & Cognition*, 48(2):226–243, 2020.

[139] Terry Connolly, Robert L Routhieaux, and Sherry K Schneider. On the effectiveness of group brainstorming: Test of one underlying cognitive mechanism. *Small Group Research*, 24(4):490–503, 1993.

[140] Nicholas W Kohn and Steven M Smith. Collaborative fixation: Effects of others' ideas on brainstorming. *Applied Cognitive Psychology*, 25(3):359–371, 2011.

[141] Andreas Fink, Roland H Grabner, Daniela Gebauer, Gernot Reishofer, Karl Koschutnig, and Franz Ebner. Enhancing creativity by means of cognitive stimulation: Evidence from an fmri study. *NeuroImage*, 52(4):1687–1695, 2010.

[142] LL Rennie and Lesley H Parker. Placing physics problems in real-life context: students' reactions and performance. *Australian Science Teachers Journal*, 42(1):55–58, 1996.

[143] 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.

[144] Stuart Glennan and Phyllis McKay Illari. *The Routledge handbook of mechanisms and mechanical philosophy*. Routledge London, UK, 2017.

[145] Michael Waldmann. *The Oxford handbook of causal reasoning*. Oxford University Press, 2017.

[146] Samuel GB Johnson and Woo-kyoung Ahn. Causal mechanisms. "", 2017.

[147] David R Shanks and Anthony Dickinson. Associative accounts of causality judgment. In *Psychology of learning and motivation*, volume 21, pages 229–261. Elsevier, 1988.

[148] Mike E Le Pelley, Oren Griffiths, and Tom Beesley. Associative accounts of causal cognition. *The Oxford handbook of causal reasoning*, pages 13–28, 2017.

[149] Leonard Talmy. Force dynamics in language and cognition. *Cognitive science*, 12(1): 49–100, 1988.

[150] Phillip Wolff. Representing causation. *Journal of experimental psychology: General*, 136(1):82, 2007.

[151] Rom Harré. Causal powers: A theory of natural necessity. "", 1975.

[152] Peter A White. A theory of causal processing. *British Journal of Psychology*, 80(4): 431–454, 1989.

[153] Lawrence W Barsalou. Perceptual symbol systems. *Behavioral and brain sciences*, 22 (4):577–660, 1999.

[154] Eugenia Goldvarg and Philip N Johnson-Laird. Naive causality: A mental model theory of causal meaning and reasoning. *Cognitive science*, 25(4):565–610, 2001.

[155] PN Johnson-Laird and S Khemlani. Mental models and causation. *Oxford handbook of causal reasoning*, pages 1–42, 2017.

[156] Daniel L Schwartz and John B Black. Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive science*, 20(4):457–497, 1996.

[157] Leonid Rozenblit and Frank Keil. The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive science*, 26(5):521–562, 2002.

[158] Thomas L Griffiths and Joshua B Tenenbaum. Theory-based causal induction. *Psychological review*, 116(4):661, 2009.

[159] Judea Pearl et al. Models, reasoning and inference. *Cambridge, UK: CambridgeUniversityPress*, 19(2), 2000.

[160] Benjamin Margolin Rottman. The acquisition and use of causal structure knowledge. *The Oxford handbook of causal reasoning*, pages 85–114, 2017.

[161] Roger C Schank and Robert P Abelson. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Psychology Press, 2013.

[162] David Hammer. Student inquiry in a physics class discussion. *Cognition and Instruction*, 13(3):401–430, 1995.

[163] Loucas T Louca, Zacharias C Zacharia, and Constantinos P Constantinou. In quest of productive modeling-based learning discourse in elementary school science. *Journal of Research in Science Teaching*, 48(8):919–951, 2011.

[164] Peter Machamer, Lindley Darden, and Carl F Craver. Thinking about mechanisms. *Philosophy of science*, 67(1):1–25, 2000.

[165] Vanessa De Andrade, Yael Shwartz, Sofia Freire, and Mónica Baptista. Students' mechanistic reasoning in practice: Enabling functions of drawing, gestures and talk. *Science Education*, 106(1):199–225, 2022.

[166] Rayendra Wahyu Bachtiar, Ralph FG Meulenbroeks, and Wouter R van Joolingen. Stimulating mechanistic reasoning in physics using student-constructed stop-motion animations. *Journal of Science Education and Technology*, 30(6):777–790, 2021.

[167] Martina A Rau. Cognitive and socio-cultural theories on competencies and practices involved in learning with multiple external representations. In *Handbook of Learning from Multiple Representations and Perspectives*, pages 17–32. Routledge, 2020.

[168] Peggy Van Meteris, Alexandra List, Doug Lombardi, and Panayiota Kendeou. *Handbook of learning from multiple representations and perspectives.* "", 2020.

[169] Elizabeth Gire and Edward Price. Structural features of algebraic quantum notations. *Physical Review Special Topics-Physics Education Research*, 11(2):020109, 2015.

[170] Ana Susac, Andreja Bubic, Petra Martinjak, Maja Planinic, and Marijan Palmovic. Graphical representations of data improve student understanding of measurement and uncertainty: An eye-tracking study. *Physical Review Physics Education Research*, 13 (2):020125, 2017.

[171] Shaaron Ainsworth. The functions of multiple representations. *Computers & education*, 33(2-3):131–152, 1999.

[172] Richard Cox. Representation construction, externalised cognition and individual differences. *Learning and instruction*, 9(4):343–363, 1999.

[173] Shaaron Ainsworth. Deft: A conceptual framework for considering learning with multiple representations. *Learning and instruction*, 16(3):183–198, 2006.

[174] Richard E Mayer. Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning*, 41:31–48, 2005.

[175] Mieke De Cock. Representation use and strategy choice in physics problem solving. *Physical Review Special Topics-Physics Education Research*, 8(2):020117, 2012.

[176] Noah S Podolefsky and Noah D Finkelstein. Use of analogy in learning physics: The role of representations. *Physical Review Special Topics-Physics Education Research*, 2 (2):020101, 2006.

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

[178] Tina Seufert. Supporting coherence formation in learning from multiple representations. *Learning and instruction*, 13(2):227–237, 2003.

[179] Leonardo Hsu, Eric Brewe, Thomas M Foster, and Kathleen A Harper. Resource letter rps-1: Research in problem solving. *American journal of physics*, 72(9):1147–1156, 2004.

[180] Christopher Pincock. A revealing flaw in colyvan's indispensability argument. *Philosophy of Science*, 71(1):61–79, 2004.

[181] Otávio Bueno and Mark Colyvan. An inferential conception of the application of mathematics. *Noûs*, 45(2):345–374, 2011.

[182] Abraham Arcavi. Symbol sense: Informal sense-making in formal mathematics. *For the learning of Mathematics*, 14(3):24–35, 1994.

[183] Patricia Heller and Mark Hollabaugh. Teaching problem solving through cooperative grouping. part 2: Designing problems and structuring groups. *American journal of Physics*, 60(7):637–644, 1992.

[184] Patrick Löffler. *Modellanwendung in Problemlöseaufgaben-wie wirkt Kontext?*, volume 205. Logos Verlag Berlin GmbH, 2016.

[185] Dara Ramalingam, Ray Philpot, and Barry Mccrae. The pisa 2012 assessment of problem solving. *""*, 2017.

[186] Argenta Price, Shima Salehi, Eric Burkholder, Candice Kim, Virginia Isava, Michael Flynn, and Carl Wieman. An accurate and practical method for assessing science and engineering problem-solving expertise. *International Journal of Science Education*, pages 1–24, 2022.

[187] Marcela Pozas, Patrick Löffler, Wolfgang Schnotz, and Alexander Kauertz. The effects of context-based problem-solving tasks on students' interest and metacognitive experiences. *Open Education Studies*, 2(1):112–125, 2020.

[188] Jongwon Park and Limook Lee. Analysing cognitive or non-cognitive factors involved in the process of physics problem-solving in an everyday context. *International Journal of Science Education*, 26(13):1577–1595, 2004.

[189] Gita Taasoobshirazi and Martha Carr. A review and critique of context-based physics instruction and assessment. *Educational Research Review*, 3(2):155–167, 2008.

[190] Judith Bennett, Fred Lubben, and Sylvia Hogarth. Bringing science to life: A synthesis of the research evidence on the effects of context-based and sts approaches to science teaching. *Science education*, 91(3):347–370, 2007.

[191] Robert Eugene Yager. *Exemplary science: Best practices in professional development.* NSTA Press, 2005.

[192] Anton Rayner. Reflections on context-based science teaching: A case study of physics for students of physiotherapy. In *Proceedings of The Australian Conference on Science and Mathematics Education*, volume 11, 2005.

[193] Nell K Duke, Victoria Purcell-Gates, Leigh A Hall, and Cathy Tower. Authentic literacy activities for developing comprehension and writing. *The Reading Teacher*, 60 (4):344–355, 2006.

[194] N İ Kurbanoğlu and Fatma Koç Nefes. Effect of context-based questions on secondary school students' test anxiety and science attitude. *Journal of Baltic Science Education*, 14(2):216, 2015.

[195] CA Ogilvie. Changes in students' problem-solving strategies in a course that includes context-rich, multifaceted problems. *Physical Review Special Topics-Physics Education Research*, 5(2):020102, 2009.

[196] Steffen P Schmidgall, Alexander Eitel, and Katharina Scheiter. Why do learners who draw perform well? investigating the role of visualization, generation and externalization in learner-generated drawing. *Learning and Instruction*, 60:138–153, 2019.

[197] Ana Susac, Andreja Bubic, Maja Planinic, Marko Movre, and Marijan Palmovic. Role of diagrams in problem solving: An evaluation of eye-tracking parameters as a measure of visual attention. *Physical Review Physics Education Research*, 15(1):013101, 2019.

[198] Michael Vignal and Bethany R Wilcox. Physics problem-solvers on why they generate unprompted diagrams. In *Proceedings of the Physics Education Research Conference (PERC*, pages 420–425, 2021.

[199] Richard Cox and Paul Brna. Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial intelligence in Education*, 6:239–302, 1995.

[200] Michael Vignal and Bethany R Wilcox. Investigating unprompted and prompted diagrams generated by physics majors during problem solving. *Physical Review Physics Education Research*, 18(1):010104, 2022.

[201] Michael Vignal and Bethany R. Wilcox. Conference proceedings detail page physics problem-solvers on why they generate unprompted diagrams. In *Physics Education Research Proceedings*, pages 396–401, 2021.

[202] 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):010114, 2020.

[203] MacKenzie Lenz, Kelby T Hahn, Paul J Emigh, and Elizabeth Gire. Student perspective of and experience with sense-making: A case study. In *PERC Proceedings*, 2017.

[204] Ronald N Giere. Using models to represent reality. In *Model-based reasoning in scientific discovery*, pages 41–57. Springer, 1999.

[205] Anjan Chakravartty. Informational versus functional theories of scientific representation. *Synthese*, 172(2):197–213, 2010.

[206] Julia Gouvea and Cynthia Passmore. Models of'versus 'models for. *Science & Education*, 26(1):49–63, 2017.

[207] Mauricio Suárez. An inferential conception of scientific representation. *Philosophy of science*, 71(5):767–779, 2004.

[208] Philip Nicholas Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness.* Harvard University Press, 1983.

[209] Niels Skovgaard-Olsen, Peter Collins, Karolina Krzyżanowska, Ulrike Hahn, and Karl Christoph Klauer. Cancellation, negation, and rejection. *Cognitive Psychology*, 108:42–71, 2019.

[210] John W Creswell and Cheryl N Poth. *Qualitative inquiry and research design: Choosing among five approaches.* Sage publications, 2016.

[211] Nitasha Mathayas, David E Brown, Robert C Wallon, and Robb Lindgren. Representational gesturing as an epistemic tool for the development of mechanistic explanatory models. *Science Education*, 103(4):1047–1079, 2019.

[212] Donald A Norman. *Some observations on mental models.* Psychology Press, 2014.

[213] Susan Goldin-Meadow and Susan M Wagner. How our hands help us learn. *Trends in cognitive sciences*, 9(5):234–241, 2005.

[214] Robert A Cortes, Adam B Weinberger, Griffin A Colaizzi, Grace Porter, Emily Dyke, Holly Keaton, Dakota L Walker, and Adam E Green. What makes mental modeling difficult? normative data for the multidimensional relational reasoning task (mrrt). *Frontiers in psychology*, 12:1512, 2021.

[215] Maxwell J Roberts. Strategies in relational inference. *Thinking & Reasoning*, 6(1): 1–26, 2000.

[216] Paul J Emigh, Jonathan W Alfson, and Elizabeth Gire. Student sensemaking about equipotential graphs. In *Physics Education Research Conference Proceedings*, 2018.

[217] David Maloney. Teaching critical thinking: Sense-making, explanations, language, and habits. *The Physics Teacher*, 53(7):409–411, 2015.

[218] Marcos D Caballero, Bethany R Wilcox, Leanne Doughty, and Steven J Pollock. Unpacking students' use of mathematics in upper-division physics: where do we go from here? *European Journal of Physics*, 36(6):065004, 2015.

[219] Ruth Boelens, Bram De Wever, and Susan McKenney. Conjecture mapping to support vocationally educated adult learners in open-ended tasks. *Journal of the learning sciences*, 29(3):430–470, 2020.

[220] Helen Wozniak. Conjecture mapping to optimize the educational design research process. *Australasian Journal of Educational Technology*, 31(5), 2015.

[221] Andrew A Tawfik, Matthew Schmidt, and Chelsy P Hooper. Role of conjecture mapping in applying a game-based strategy towards a case library: a view from educational design research. *Journal of Computing in Higher Education*, 32(3):655–681, 2020.

[222] Victor R Lee, Mimi Recker, and Abigail L Phillips. Conjecture mapping the library: Iterative refinements toward supporting maker learning activities in small community spaces. *""*, 2018.

[223] Linda E Strubbe, Adrian M Madsen, Sarah B McKagan, and Eleanor C Sayre. Beyond teaching methods: Highlighting physics faculty's strengths and agency. *Physical Review Physics Education Research*, 16(2):020105, 2020.

[224] Jason M May, Lauren A Barth-Cohen, Jordan M Gerton, Claudia De Grandi, and Adrian L Adams. Student sensemaking about inconsistencies in a reform-based introductory physics lab. *Physical Review Physics Education Research*, 18(2):020134, 2022.

[225] Meagan Sundstrom, Anna McLean Phillips, and NG Holmes. Problematizing in inquiry-based labs: how students respond to unexpected results. In *Proceedings of the 2020 physics education research conference, virtual conference*, 2020.

[226] Laura McCullough. Gender, context, and physics assessment. *Journal of International Women's Studies*, 5(4):20–30, 2004.

[227] Stephen K Reed. The structure of ill-structured (and well-structured) problems revisited. *Educational Psychology Review*, 28(4):691–716, 2016.

[228] Namsoo Shin, David H Jonassen, and Steven McGee. Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of research in science teaching*, 40(1):6–33, 2003.

[229] Javier Alejandro Pulgar. *Performance on Well and Ill-Structured Problems in University Physics: The Role of Instruction in Cooperative Learning*. University of California, Santa Barbara, 2019.

[230] Javier Pulgar, Cristian Candia, and Paul M Leonardi. Social networks and academic performance in physics: Undergraduate cooperation enhances ill-structured problem elaboration and inhibits well-structured problem solving. *Physical Review Physics Education Research*, 16(1):010137, 2020.

[231] Victor Law, Xun Ge, and Kun Huang. Understanding learners' challenges and scaffolding their ill-structured problem solving in a technology-supported self-regulated

learning environment. In *Handbook of research in educational communications and technology*, pages 321–343. Springer, 2020.

[232] Ying Chen, Paul W Irving, and Eleanor C Sayre. Epistemic game for answer making in learning about hydrostatics. *Physical Review Special Topics-Physics Education Research*, 9(1):010108, 2013.

Appendix A

A.1 Scientific practices

(A) Developing and Using Models.

(B) Planning Investigations.

(C) Analyzing and Interpreting Data.

(D) Using Mathematics and Computational Thinking.

(E) Constructing Explanations and Engaging in Argument from Evidence.

(F) Evaluating Information.

A.2 Cross-cutting concepts

(a) Patterns.

(b) Cause and Effect: Mechanism and Explanation.

(c) Scale and proportion.

(d) Proportion and Quantity.

(e) System and System Models.

(f) Energy and Matter: Flows, Cycles, and Conservation.

(g) Structure and Function.

(h) Stability and Change.

A.3 Sub-ideas

(i) Energy is conserved.

(ii) Temperature can be defined in terms of changes in entropy and internal energy.

(iii) Spontaneous processes occur such that entropy is maximized.

(iv) The macroscopic quantity of entropy for a system can be determined using the number of microstates within that system.

(v) Thermodynamic potentials can be used to determine the energy that is available to perform work.

(vi) Heat flow changes (or does not change) temperature depending on the process undergone and ambient conditions

(vii) The macroscopic quantity of entropy for a system can be determined using the number of microstates within that system.

(viii) Probability.

(ix) Microstates.

(x) States/Boltzmann Factor.

A.4 Learning performances

1. **(E-f-i)** Construct an argument justifying or refuting claims about the changes to internal energy of a thermodynamic system given information about the energy flow into and out of the system.

2. **(E-d-ii)** Construct an argument justifying or refuting claims about the temperature of a system using information about changes in entropy and internal energy.

3. **(C-b-iii)** Analyze and interpret data about interacting systems to determine whether a thermodynamic process will happen spontaneously using the idea that entropy of the universe is maximized for spontaneous processes.

4. **(D-c-iv)** Use mathematics to determine the number of microstates within a system to deduce the macroscopic quantity of entropy for that system and make a conclusion about the system.

5. **(E-f-v)** Construct an argument justifying the most appropriate thermodynamic potential (e.g., Gibbs free energy, Helmholtz free energy, enthalpy) to be applied to a given context depending on system conditions (e.g., constant pressure, volume, or temperature).

6. **(E-b-vi)** Generate an explanation about the mechanism by which the temperature does (or does not) change with heat flow into or out of a system informed by the process undergone and ambient conditions (e.g., pressure, temperature).

7. **(F-a-vii)** Evaluate information in the form of ideas generated by students about the entropy of a system undergoing changes to its state by considering the number of microstates for a given macrostate.

8. **(A-c-viii)** Use models to determine the number of microstates for a given macrostate and find the probability of a system being in that particular macrostate.

9. **(A-c-viii)** Use a model to determine the probability of a system being in a particular state using the Boltzmann factor for that particular state and make relevant conclusions.
10. **(A-c-i)** Use representation(s) of the speed distribution of a gas to reason about its composition or properties.
11. **(D-f-i)** Use mathematics to calculate the temperature and internal energy and use them to predict macroscopic features of an Einstein solid given information about the entropy.
12. **(E-f-i)** Engage in argumentation regarding how (or if) a system's partition function is impacted by contact with another system.
13. **(D-c-i)** Use mathematics to determine the heat capacity of argon and helium from the monatomic gas partition function and see if either gasses are usable in an experiment.

Appendix B

Feedback statements

1. *You wanted to assess if your students are able to:* Construct an argument justifying or refuting claims about the changes to internal energy of a thermodynamic system given information about the energy flow into and out of the system.

Students were asked to:

1. Unpack the relations between the changes in internal energy of a system with respect to heat transferred in/out of the system and work done on/by the system ($\Delta U = Q \pm W$).
2. Use the unpacked relations to generate an explanation about the changes in internal energy of the system.
3. Make a claim about the changes in internal energy of the system.

The TaSPA:
(ES1)

M	Provided evidence that your students unpacked the relations between the change in internal energy of a system with respect to heat transferred in/out of the system and work done on/by the system.
P	Provided some evidence that your students related changes in changes in internal energy of a system to either heat or work as forms of energy flow into and out of that system, but not both.
N	Did not provide evidence that your students unpacked the relations between the change in internal energy of a system with respect to heat transferred in/out of the system and work done on/by the system.

(ES2)

M	Provided evidence that your students constructed arguments about the changes to internal energy of a system by taking into account the contributions from both heat and work as forms of energy flow into and out of that system.
P	Provided some evidence that your students constructed arguments about the changes to internal energy of a system by taking into account the contributions from either heat or work as forms of energy flow into and out of that system, but not both.
N	Did not provide evidence that your students constructed arguments about the changes to internal energy of a system by taking into account the contributions from both heat and work as forms of energy flow into and out of that system.

(ES3)

Students could benefit from more opportunities to:

(ES1)

M	Provided evidence that your students made an accurate claim about the changes to internal energy of a system.
N	Did not provide evidence that your students made an accurate claim about the changes to internal energy of a system.

P	explore how both heat and work as forms of energy flowing into and out of real-world systems can relate to changes in internal energy of that system.
N	explore factors that contribute to changes in internal energy of a system using ideas of conservation of energy, such as the first law of thermodynamics.

(ES2)

P	generate clear explanations about how both heat and work as forms of energy in real-world systems can concurrently contribute to the changes in internal energy of a system.
N	generate clear explanations about changes in internal energy of a system when the concerned process involves concurrent contributions from both heat and work.

(ES3)

N	make conclusions about the overall changes to internal energy of a system based on properties of a system and/or characteristics of a particular process.

2. *You wanted to assess if your students are able to:* Construct an argument justifying or refuting claims about the temperature of a system using information about changes in entropy and internal energy.

Students were asked to:

1. extract a mathematical relationship between temperature, entropy and internal energy using the provided thermodynamic identity.
2. use the extracted mathematical relationship to determine the temperature values of the system.
3. validate the provided claim or hypothesis using the given data and the extracted mathematical relationship.

The TaSPA:

(ES1)

M	Provided evidence that your students extracted the requisite mathematical relationship between temperature, entropy and internal energy.
P	Provided some evidence that your students partially extracted the required mathematical equation by incompletely mapping the information from the physical system to the provided thermodynamic identity.
N	Did not provide evidence that your students extracted the requisite mathematical relationship between temperature, entropy and internal energy using the provided thermodynamic identity.

(ES2)

M	Provided evidence that your students used the extracted mathematical relationship to determine the temperature values.
P	Provided some evidence that your students used the given thermodynamic identity to reason about the temperature of the given system either by deducing an incorrect relation, or by not determining temperature's appropriate values.
N	Did not provide evidence that your students coordinated between the provided information on changes in entropy and internal energy of the system to reason about temperature.

(ES3)

M	Provided evidence that your students validated the provided claim or hypothesis using the provided data and the extracted mathematical relationship.
N	Did not provide evidence that your students validated the provided claim or hypothesis by coordinating the provided thermodynamic identity with information on changes in entropy and internal energy of the system.

Students could benefit from more opportunities to:

(ES1)

P	explore physical features of real-world systems and their mathematical implications while reasoning about thermodynamic identities or quantities such as temperature.
N	reason about the relationship between a system's temperature and changes in its internal energy and entropy.

(ES2)

P	interpret data on (changes in) internal energy and entropy to reason about temperature of systems.
N	explore the relationship between internal energy and entropy of the system with respect to its temperature.

(ES3)

N	validate observations about thermodynamic features such as temperature of real-world systems by leveraging contextual data and thermodynamic relationships.

3. *You wanted to assess if your students are able to:* Use mathematics to determine the number of microstates within a system to deduce the changes in macroscopic quantity of entropy for that system and make a conclusion about the system.

Students were asked to:

1. identify an accurate relationship between the number of microstates and entropy of the system.
2. calculate the number of microstates for a given macrostate of the system.
3. determine the entropy of the system.
4. interpret the obtained entropy value to make a conclusion about the system.

The TaSPA:

(ES1)

M	Provided evidence that your students identified an accurate relationship between the number of microstates and entropy of the system.
N	Did not provide evidence that your students identified an accurate relationship between the number of microstates and entropy of the system.

(ES2)

M	Provided evidence that your students appropriately calculated the number of microstates for a particular macrostate of the given system.
N	Did not provide evidence that your students calculated the number of microstates for a particular macrostate of the given system.

(ES3)

M	Provided evidence that your students accurately determined the entropy of the system.
N	Did not provide evidence that your students accurately determined the entropy of the system.

(ES4)

M	Provided evidence that your students accurately interpreted the obtained entropy to make a conclusion about the system.
N	Did not provide evidence that your students accurately interpreted the obtained entropy to make a conclusion about the system.

Students could benefit from more opportunities to:

(ES1)

N	explore how the number of accessible microstates of a given system's macrostate relate to its entropy.

(ES2)

N	reason about multiplicity, permutations, combinations, and logarithmic operations in calculating the number of microstates for a given macrostate of the system.

(ES3)

N	reason about how entropy of a system is related to the number of microstates of a given macrostate.

(ES4)

N	reason about the changes in entropy and its implications for the spontaneity of real-world thermodynamic processes.

4. *You wanted to assess if your students are able to:* Analyze and interpret data about interacting systems to determine whether a thermodynamic process will happen spontaneously using the idea that entropy of the universe is maximized for spontaneous processes.

Students were asked to:

1. identify maximization of entropy as crucial for spontaneous processes.
2. use the given data representation to determine the entropy of the system.
3. validate the given claim or hypothesis through a conclusion.

The TaSPA:

(ES1)

M	Provided evidence that your students identified maximization of the entropy of the universe as the driver in a spontaneous process.
N	Did not provide evidence that your students identified maximization of the entropy of the universe as the driver in a spontaneous process.

(ES2)

M	Provided evidence that your students used the given data representation to determine the required entropy of the system.
P	Provided some evidence that your students either extracted the necessary information from the data representation, or reasoned correctly about the changes in the entropy of the given system, but not both.
N	Did not provide evidence that your students used the provided data representation to reason about the entropy of the system.

(ES3)

M	Made an accurate conclusion about the spontaneity of the thermodynamic processes based on the maximization of the universe's entropy.
N	Did not make an accurate conclusion about the spontaneity of the thermodynamic process by accounting for the maximization of the universe's entropy.

Students could benefit from more opportunities to:

(ES1)

N	explore the relationship between entropy and spontaneity in real-world thermodynamic processes.

(ES2)

P	interact with various forms of data representations about thermodynamic quantities associated with real-world systems.
N	reason about the spontaneity of real-world thermodynamic processes based on data pertaining to various thermodynamic quantities, particularly entropy.

(ES3)

N	make conclusions about the spontaneity of everyday thermodynamic processes based on the maximization of the universe's entropy.

5. *You wanted to assess if your students are able to:* Construct an argument justifying the most appropriate thermodynamic potential (e.g., Gibbs free energy, Helmholtz free energy, enthalpy) to be applied to a given context depending on system conditions (e.g., constant pressure, volume, or temperature).

Students were asked to:

1. Identify the thermodynamic conditions of the described system.
2. justify the choice of a thermodynamic potential based on the given conditions.
3. make a claim about the appropriate thermodynamic potential applicable to the given context.

The TaSPA:

(ES1)

M	Provided evidence that your students identified the thermodynamic conditions of the described system.
N	Did not provide evidence that your students identified the thermodynamic conditions of the described system.

(ES2)

M	Provided evidence that your students justified the choice of a thermodynamic potential informed by the given conditions.
P	Provided some evidence that your students either justified the appropriate thermodynamic potential or identified the physical conditions which describe usable work for the given system.
N	Did not provide evidence that your students justified the choice of a thermodynamic potential informed by the given conditions.

(ES3)

M	Provided evidence that your students made a claim about the appropriate thermodynamic potential applicable to the given context.
N	Did not provide evidence that your students made a claim about the appropriate thermodynamic potential applicable to the given context.

Students could benefit from more opportunities to:

(ES1)

N	extract relevant information from everyday systems/processes in order to identify the thermodynamic conditions.

(ES2)

P	identify the physical quantities that determine the choice of an appropriate thermodynamic potential to describe usable work in everyday processes.
N	identify the factors that determine the choice of appropriate thermodynamic potential to describe usable work for systems/processes.

(ES3)

N	justify or refute the existing claims, particularly as they relate to thermodynamic potentials of real-world processes.

6. *You wanted to assess if your students are able to:* Generate an explanation about the mechanism by which the temperature does (or does not) change with heat flow into or out of a system informed by the process undergone and ambient conditions (e.g., pressure, temperature).

Students were asked to:

1. identify relevant process(es) undergone and ambient conditions of the system that will influence whether or not temperature changes.
2. describe the physical mechanism that justifies why heat does or does not change temperature.

The TaSPA:

(ES1)

M	Provided evidence that your students identified relevant process(es) undergone and ambient conditions of the system that will influence whether or not temperature changes.
N	Did not provide evidence that your students identified relevant process(es) undergone and ambient conditions of the system that will influence whether or not temperature changes.

(ES2)

M	Provided evidence that your students gave a complete description of a physical mechanism that justifies why heat does or does not change temperature.
P	Provided some evidence that your students partially identified a physical mechanism that justifies why heat does or does not change temperature.
N	Did not provide evidence that your students described a physical mechanism that justifies why heat does or does not change temperature.

Students could benefit from more opportunities to:

(ES1)

N	identify the relevant processes and ambient conditions which influence the changes in temperature of a system given the heat flow into and out of the system.

(ES2)

P	explore the physical mechanisms through which the heat flow into or out of the system affects its temperature.
N	examine the processes undergone and the ambient thermodynamic conditions while investigating the temperature fluctuations in real-world systems.

7. *You wanted to assess if your students are able to:* Evaluate information in the form of ideas generated by students about the entropy of a system undergoing changes to its state by considering the number of microstates for a given macrostate.

Students were asked to:

1. Identify the relation between entropy and the number of microstates for a given macrostate of the system.
2. determine the change in the number of microstates for the given macrostate of the system.
3. validate the provided information about the entropy of the system through the unpacked relation.

The TaSPA:

(ES1)

M	Provided evidence that your students identified the relation between entropy and the number of microstates for a given macrostate of the system.
N	Did not provide evidence that your students identified the relation between entropy and the number of microstates for a given macrostate of the system.

(ES2)

M	Provided evidence that your students determined the change in the number of microstates for the different macrostates of the system.
P	Provided some evidence that your students partially reasoned about the change in the number of microstates for the specified macrostate of the physical system.
N	Did not provide evidence that your students determine the change in the number of microstates for the given macrostate of the system.

(ES3)

M	Provided evidence that your students validated the provided information about the entropy of the system.
N	Did not provide evidence that your students validate the provided information about the entropy of the system through the unpacked relation.

Students could benefit from more opportunities to:

(ES1)

N	explore how microscopic properties determine the macroscopic features of a physical system.

(ES2)

P	explore how the variation in the system's macroscopic parameters lead to the changes in the number of microstates for the corresponding macrostate of the system.
N	explore how the number of microstates is associated with a macrostate of real-world systems by employing mathematical relations such as Sackur-Tetrode equation and principles such as multiplicity, permutations, combinations, and logarithmic operations.

(ES3)

N	validate the provided information on changes in entropy of a system by taking into account the corresponding changes in the number of microstates for a given macrostate.