An Empirical Study of the Process of Crafting and Using Definitions

By

Angela Jean Little

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Science And Mathematics Education in the Graduate Division of the University of California, Berkeley

Committee in charge:

Professor Andrea A. diSessa, Chair
Professor Alan H. Schoenfeld
Professor Eve E. Sweetser

Fall 2013
An Empirical Study of the Process of Crafting and Using Definitions

Copyright, 2013

By

Angela Jean Little
Abstract

An Empirical Study of the Process of Crafting and Using Definitions

by

Angela Jean Little

Doctor of Philosophy in The Graduate Group in Science and Mathematics Education

University of California, Berkeley

Professor Andrea diSessa, Chair

In this dissertation I analyze the process of crafting definitions whose purpose is classification. The context I examine is undergraduate upper-division physical science majors defining and naming sub-categories of a physical phenomenon in the context of a design task over an extended period of time. The goal of the design task is one of classification: help people better identify the phenomenon out in the world.

I first develop an analytic framework for the process of improving a definition. This framework involves an interplay between four main elements: the current state of the definition, criteria for what makes a good definition, examples, and definitionally unarticulated knowledge (DUK). By DUK, I mean implicit judgments that definers make about categorization that have not been incorporated into the definition, per se. I show how criteria and the practices associated with meeting them guide the refinement work.

Once participants craft their definition a question emerges: to what extent does it serve them as a tool when they make sense of new candidate examples? I zoom in on moments where new candidate examples are introduced by participants or the facilitator and describe the extent to which students rely on their own crafted definition. The consideration of a new candidate example can sometimes lead to participants naming a new subcategory and I analyze this process as well.

The detailed empirical analysis is bookended by a set of workshop design principles that made the analysis possible and some educational implications of the work. I lay out a set of design principles for creating activities wherein people engage productively in collaboratively defining. I conclude by exploring the educational implications for college physics teaching as well as more general instances where one is interested in crafting a definition for the purpose of classification.
Acknowledgements

I have to thank my parents first and foremost. Their encouragement and emotional support all throughout my K-12, college, and graduate schooling were central to my finishing this Ph.D. Others that played particularly pivotal mentorship roles along the way were my high school physics teacher, Michael Lampert, and my undergraduate honors thesis advisor, Stamatis Vokos. Without Michael suggesting that I major in physics and without the incredible community and support that I found in SPU’s physics department, including Stamatis’ wonderful mentorship and support in particular, I would not be where I am today. I also am deeply grateful for all of the graduate and undergraduate students involved in The Compass Project, a program I co-founded during my time as a Ph.D. student in the physics department at Berkeley. I discovered a calling in Compass and developed friendships that supported me when I thought I would leave graduate school. Compass gives me hope for positive change in the world.

Before I thank them for their intellectual support of my research, I must first thank my committee members for their mentorship during a difficult decision I had to make about switching from my physics Ph.D. program to an interdisciplinary program. Eve Sweetser was first in this: I took her endlessly interesting class The Mind and Language during a semester when I was beginning to understand that I had more interdisciplinary interests but was still a physics graduate student. We discussed the possibility of my being interdisciplinary between the linguistics and physics departments. It turned out that an interdisciplinary graduate group already existed in STEM education, however. Around the same time, I happened to be hired to co-design and co-teach a course with Alan Schoenfeld and Dor Abrahamson. I may not recall the exact specifics (was it Alan or Dor who first suggested I consider SESAME?), but Alan was instrumental in supporting me to switch over and mentoring my early education research. He introduced me to Andy diSessa who empathized with my physics to STEM education transition and turned into my primary advisor as I finished.

Andy was also incredibly supportive when I had difficult injuries from a bicycle accident, helping me get back on my feet by meeting with me weekly for a couple of months when I’d finally healed. Thank you to all of my aunts and uncles and grandmother (rest in peace, Lucy) who banded together and supported me financially when I would have gone into debt from the dental bills associated with the accident. It helped me stay focused on my dissertation.

Intellectually, I am thankful for a committee that has always treated me like a colleague. I really can’t tell you how much I am appreciative of that. Every step of the way, their feedback has felt much less like harsh judgment and much more like thoughtful experts who want the best for my burgeoning work. I think I will forever more have Andy’s voice in my head when I’m editing papers that make general sweeping statements (including my own): “Do you really mean EVERYTHING works like that? I bet I can find a counter example!” I’ve learned so much about all aspects of writing and researching from Andy. From Alan, I have to say that foremost I think about his incredible skill of top-level organization. I’ve so
often found myself in his office tossing out a set of mushy ideas and then getting to watch as he organizes it into a coherent flow of ideas that make so much more sense. It's such an incredibly useful skill that has helped me improve my work a ton. From Eve, I picked up my fascination with and close attention to language. She showed me how much comes out of seemingly subtle differences in language.

In addition to my committee, SESAME and EMST have an incredible group of graduate students, many of whom I count as friends as well as colleagues. A huge thank you to the Patterns and Functions research groups where I received incredible feedback on my work and found a warm and supportive community that cared so deeply about students and STEM education. Thank you to each and every member. In particular, I don’t think I would have survived my qualifying exam studying without Kim, who made it extra special with her lobster cookies. Hillary: I’m so glad we could talk about work in the context of life goals more generally and share strategies for holistic living. Lauren and Colleen, what better officemates could a lady ask for? Thank you to people who provided critical last minute feedback on aspects of my dissertation: Gina, Lama, Lauren, and Colleen. Thank you to Kate Capps, also, our SESAME administrative support person who listened, cared deeply about us, shepherded us through all of the bureaucratic requirements. Although she isn’t “SESAME,” I should also thank Anne Takizawa, our administrative support person for graduate students in the physics department. She was one of my biggest supporters during my three years in the physics department and I was so glad I had her as a listening ear. Plus, she made me applesauce when I had my bike accident and couldn’t eat solid foods.

There were other important supports to my dissertation: my counselor Lisa, my friend Denia who, what can I say, we “ran the race together,” my partner Mel who drove me to the coffee shop on cold mornings when I lacked motivation to write and helped me with annoying Microsoft Word format issues near the end. My friend Jess, who listened to me whenever I needed it that final semester. I am incredibly thankful for two coffee shops, their warm staff members, and tasty coffee: The Coffee Studio in Chicago, IL, and Local 123 in Berkeley, CA. Thank you for being a positive space where I looked forward to writing and for all the cappuccinos and “physics-ccinos” that got me through this.

Well, I’m starting to tear up now because you see how my finishing my Ph.D. was due to the support of so many incredible people along the way. I mean, seriously. It’s ridiculous. Truly, I couldn’t have done it without them. I hope to live a life of supporting others because they have supported me.

I would like to thank some additional people that did not fit neatly into any of the previous paragraphs. Kathleen Metz for all of the wonderful discussions about qualitative research methodology both in her course and as a member of my qualifying exam committee. Also, I learned a great deal and had way too much fun in Coye Cheshire’s quantitative methods course. Best course ever! Marjorie Shapiro and Bob Jacobsen for their support of my educational interdisciplinary interests in the physics department. Nicci Nunes for hiring me to teach for Cal Teach. It was an incredible learning experience and influenced my switch to SESAME. Nicci herself provided so much wonderful mentorship. Also Dor
Abrahamson who I co-taught the Cal Teach course with for three years and learned a great deal from. And Nilofar, for all of the support and friendship during graduate school, particularly during my qualifying exam preparation (yay for de-stressing with David Attenborough documentaries about birds). And Kristin, for all the fun swimmuritos and le med dinners and encouraging notes along the way. And all of my wonderful Cal Teach, CS301, physics, and Compass students who inspired me and who I learned a great deal from. And Mari and Eric who co-thought with me about my dissertation title.

Finally, I thank all of my dear friends and family. I’m so thankful for you all for many reasons, but, pertaining to this dissertation, your support really got me through graduate school.
Table of Contents

Chapter 1: Preface and Literature Review ................................................................. 1
  Personal Preface: Background in Designing Curriculum on Crafting Definitions .... 1
  Literature Review .................................................................................................... 1
    Existing Research On Defining ........................................................................... 1
    Purposes of Crafting Definitions ........................................................................ 4
    Research Questions ............................................................................................ 4
    Related Research: Influences on Classification and the Context Dependent Nature of Knowledge ................................................................. 4
    Related Research: The Process of Crafting Representations ............................. 5

  Dissertation Overview .......................................................................................... 6

Chapter 2 – Workshop Design and Data Collection Methodology ..................... 8
  Summary of Chapter ............................................................................................. 8
    Summary of Workshops and Description of Threshold ...................................... 8
      What is Threshold? What contributes to people’s experience with the phenomenon? 8
    Overview of Workshops ................................................................................... 9

  Design Goals and Design Choices ...................................................................... 10
    Design Variable Choices .................................................................................... 11

  Workshop and Interviewer Script ....................................................................... 16
    Introductory Script For Both Small Group and Individual Participants: .......... 17
    Script for Interaction After Participants Craft Definition ............................... 20

  Artifacts Collected ............................................................................................. 22

  Room Context and Video Set-up .......................................................................... 22

  Transcript Convention ......................................................................................... 23

  Conclusion .......................................................................................................... 23

Chapter 3 – Three Episodes Of Definitional Refinement: Identification of Criteria and Practices Associated with Meeting Them ................................. 24

  Introduction ........................................................................................................ 24

  Analysis Methods: Development of Analytic Focus and Creation of Analytic Framework for Definitional Refinement ..................................................... 24
    Development of Initial Framework ................................................................ 24
    Development of an Improved Analytic Framework ....................................... 25
    Adding A New Framework Element ................................................................ 26
    Final Framework For The Process of Definitional Refinement ....................... 27
    Better Criteria Identification .......................................................................... 28
    Identification of ‘Practices Associated with Meeting Criteria’ ....................... 28

  Data and Analysis ............................................................................................... 28
    Overview ........................................................................................................... 28

    Definitional Refinement: Episode One .......................................................... 29
    Definitional Refinement: Episode Two ......................................................... 34
    Definitional Refinement: Episode Three ....................................................... 39

    Conclusion ....................................................................................................... 44

Chapter 4 – A Cross-Case Analysis of How Participants Make Sense of New candidate examples ....................................................................................... 47

  Introduction ........................................................................................................ 47

  Methodological Details about DUK ................................................................. 48
    Identifying Use of Crafted Definition in Transcript Data ............................... 48
Chapter 1: Preface and Literature Review

Personal Preface: Background in Designing Curriculum on Crafting Definitions
This dissertation is about the process of crafting and using definitions. Although I focus on particular phenomena and contexts that are physics related, my goal is to build a general framework for the process. I will also examine the benefits of educators involving their students in crafting definitions when I address the educational implications of my work.

Defining is an incredibly rich process that plays out in a wide range of contexts. In graduate school, my teaching has frequently involved supporting my students in crafting conceptually central definitions. I developed and taught curriculum where freshman physics students worked to define mechanical wave in the context of addressing a research question about what earthquakes can tell us about the interior of the earth. I co-led a discussion where a class of freshman in physics had to come to consensus on the definition of physics model to support their work in using models in pursuing their own research questions. Defining happened in more than just the physics classes I developed. I facilitated workshops for instructors who had to co-design curriculum together. They thus needed a set of agreed-upon principles to appeal to during design decisions. The process of articulating these principles was a definitional conversation: we asked everyone to share out examples of educational experiences that had been positive and negative. We then drew out themes and used this as a starting point to refine a set of articulated principles for what we wanted to achieve in the classroom. You could call it a definition for ideal classroom. I’ve recently led workshops on defining proudness where participants share out examples of things they’ve been proud of and develop a set of principles to create educational experiences where their students can be proud of their work. As an educational researcher working on the process of defining, I have even taken on the meta-task of defining defining, even. The process of crafting a definition is an interesting and challenging design task. My background in designing and facilitating definitional discussions heavily informs the research described here. In addition to building an analytic framework for the process of crafting definitions, I will speak to some design principles for building productive definitional conversations.

Literature Review

Existing Research On Defining
The process of crafting definitions is an important yet understudied part of professional practice. STEM classrooms that take an inquiry approach and attempt to bring students into authentic scientific or mathematical practice often engage in crafting definitions. Researchers have examined classroom conversations where students define terms like alive (Warren and Roseberry 2011; Bang, et al. 2013) and whether orcas are whales or dolphins (Engle and Conant, 2002 ). In these cases, definitional conversations underlie framework building for topics like productive disciplinary engagement and how desettling can support the inclusion of epistemological commitments of non-dominant communities. As for research
examining the process of crafting definitions, however, there is surprisingly little in science education, though there is some in mathematics education.

One of the earliest studies to describe the process of crafting definitions in mathematics is Fawcett's book (Fawcett, 1938) on teaching a high school geometry course that took a different approach from the more traditional textbook problems of the time. Instead, Fawcett supported his students in the practices of professional mathematicians, which included crafting and defending definitions. Schoenfeld, in his book *Learning To Think Mathematically*, gives a thorough summary (Schoenfeld, 1992):

> Fawcett’s goals were that students develop a good understanding of the subject matter of geometry, the right epistemological sense about the mathematics, and a sense of the applicability of the reasoning procedures that they had learned in geometry to situations outside of the mathematics classroom. In order for this to happen, he believed, (1) the students had to engage in doing mathematics in a way consistent with his mathematical epistemology, (2) the connections between mathematical reasoning in the formal context of the classroom and mathematical reasoning outside of it would have to be made explicit, and (3) the students would need to reflect both on their doing of mathematics and on the connections between the reasoning in both contexts...

> Fawcett pointed out that definitions have consequences: in his school, for example, there was an award for the "best teacher." Many students favored the librarian -- but was the librarian a teacher? Or, he used sports as an analogy. In baseball, for example, there might be varying definitions of "foul ball" (is a fly ball that hits the foul pole fair or foul?) -- but once one sets the rules, the game can be played with consistency. After such discussions, Fawcett notes "[n]o difficulty was met in leading the pupils to realize that these rules were nothing more than agreements which a group of interested people had made and that they implied certain conclusions" (p.33). In the mathematical domain, he had his students debate the nature and usefulness of various definitions. Rather than provide the definition of "adjacent angle," for example, he asked the class to propose and defend various definitions.

Fawcett’s work was some of the earliest to suggest that it was important to engage students in doing what mathematicians do and also show that it was possible to do so.

Contemporary research that focuses on crafting definitions generally falls into two camps: existence proofs that students can productively engage in crafting definitions for important terms over weeks-long curriculum (Atkins and Salter, 2010; Swinyard, 2011) and more general framework building (Nachlieli and Sfard, 2003; Zandieh and Rasmussen 2010). These two more general frameworks focus on opposite ends of the spectrum with respect to the length of time that people engage in defining. Nachlieli and Sfard’s work on “the activity of defining” (AoD) is focused on shorter clarifying moves. They define AoD as “the activity that aims at keeping communication effective by clarifying the use of words” and note some important practices associated with this that they name “meta-discursive” practices.
For instance, they provide data in a mathematics classroom where one student asks another “what is a cone?” and a clown’s hat that is in the room is pointed to. This pointing to a particular example, i.e. ostensibly defining, is one effective metadiscursive practice discussed. Nachlieli and Sfard choose to focus on clarifying moves and do not examine an extended process of crafting a definition.

Zandieh and Rasmussen, however, focus on an extended multi-week curriculum (Zandieh and Rasmussen, 2010). They are interested in “the role that defining can play in students’ transition from less formal to more formal ways of reasoning” and build a “Defining as a Mathematical Activity” (DMA) framework. This framework, “brings together the four levels of activity from [Realistic Mathematics Education] with the notions of creating and using concept images and concept definitions.” Concept image and concept definition are two constructs from Tall and Vinner’s work in looking at whether and how students use formal mathematics definitions (Tall and Vinner, 1981). Concept image refers to, “the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes.” Concept definition is “the words used to specify that concept” which can be memorized or more meaningfully learned. They also specify that someone’s “personal concept definition” might be different from the “formal concept definition” which is “accepted by the mathematical community at large.” Use of concept image/concept definition suggests that one target of instruction is to bring the concept image in line with the formal concept definition if there exists an agreed-upon one. It also problematizes the practice of giving students definitions to use in math class: just because a student has memorized a definition does not mean that he or she will be able to or choose to use it.

In this dissertation I will also build a general framework for crafting definitions. This framework differs from the mathematics education community’s frameworks in two ways. First, with respect to the length of the defining activity, my framework falls between AoD and DMA: I examine an hours-long extended process in depth. This allows me to say more about when and how definitional refinement occurs and what is important to it. Another important difference is my subject matter: general framework building for defining has traditionally focused on mathematics. My framework will be built from data of physical-science college students crafting a definition of a more science-related word.

There exists one other study in physics education research that examines an hour-long defining conversation in depth. This is Quan’s interaction analysis of a whole-class consensus conversation focused on how to define physics model with physical science freshman (Quan, 2012; Quan and Little, 2012). It is perhaps not surprising that Quan’s work is related to my own since she is a close colleague and her analysis focused on video data from a classroom discussion that I was co-leading. Quan builds a coding scheme to characterize consensus moves. In doing so, she examines how a group of students rejected or accepted particular claims. My framework building occurs at a different grain size: I name elements important to definitional refinement and ask what practices undergraduates bring to this process. Additionally, I focus on small group conversations rather than whole-class discussions.
Purposes of Crafting Definitions
Definitions can be designed for different purposes. For my framework building I will focus on the purpose of classifying. In science and science education, definitions are often further refined or even initially crafted in the context of experimental measurement and modeling activities (such as bluriness in Atkins and Salter’s work). My framework might prove useful in understanding definitional activities whose purpose is supporting experimental measurement and modeling, but I will leave this as an open empirical question.

Definitions take on particular classificational significance in the areas like law and mathematics. Whether a piece of land is considered a wetland, for instance, has considerable affect on how that land is maintained and used according to U.S. policy (Schiappa, 1999). In mathematics, whether a particular mathematical object counts as a polyhedron can be the difference between whether or not a mathematician can complete a proof (Lakatos, 1976). Definitions of physical phenomena can make an impact on what contexts students see as connected to one another.

Research Questions
Given the importance of the process of crafting definitions toward supporting classification, what does this process look like? What is important to it? And, once someone crafts a definition, what work does it actually do for her when she considers a new candidate example with it? These are the central research questions to my work.

Related Research: Influences on Classification and the Context Dependent Nature of Knowledge
Research in cognitive science suggests that our experience of being exposed to certain kinds of objects or contexts more than others can greatly influence categorization and reasoning. For instance, in studying the category of bird, Rips found that research subjects had category gradations related to typicality that influenced them to reason asymmetrically about the spread of bird disease: that it more easily spread outward from central members (Rips, 1975).

Research on schemas, scripts, and frames (for a summary see Lakoff, 1987; on frames see Minsky, 1975; Tannen, 1993) suggests that we develop certain expectations and assumptions around events and interactions that help with our cognitive processing. These assumptions can have an impact on our reasoning as well. For instance, our exposure to popular media and experience in hospitals with majority male surgeons can give rise to our being confounded by a popular riddle where we must imagine that a surgeon could be female to solve it.

Research in physics education on student understanding of concepts like force shows that our ability to “see” force in a particular scenario can be strongly context dependent (diSessa, 1993; diSessa, 2006; diSessa, 2008). One theory of conceptual change, Coordination Class Theory, discusses learning as a student increasing the span and alignment of readout strategies in considering what force

---

1 The full riddle: A father and his son are in a car accident. The father dies at the scene and the son is rushed to the hospital. At the hospital the surgeon looks at the boy and says “I can't operate on this boy, he is my son.” How can this be?
might be present across and within many different contexts (diSessa and Sherin, 1998). The authors define alignment as “making sure that, in different situations, the same information is read out” and span as “covering readout in the breadth of situations necessary.” Another way to say this is that learning involves identifying and honing the best information to pay attention to in the world, particularly in contexts that might throw off our initial intuition.

Cognitive linguistics and physics education research suggest that one must be careful in understanding the tacit influences of perceptual exposure on how we categorize. They also suggest that subtle changes in context or framing have big influences on what we “see.” This raises a question about definitions: if definitions are supposed to help us classify, then to what extent do they help us uncover or override these tacit influences on classification?

Research in law education suggests that the interplay between a definition and a set of related cases or “hypos” (hypothetical cases) can help uncover tacit influences on classification (Rissland, 1983) and is critical to the process of refining a definition. Rissland discusses the general approach of a law class: “A proposition or doctrinal position is enunciated and then explored by considering it on a sequence of cases. The proposition is then usually refined and the process repeats itself.” She also describes how the sequence of cases is generated. She examines a particular class working on understanding strong and weak intentionality. The professor starts with a simplified “seed” case of a lawsuit involving one kid kicking another kid at school. The class then makes modifications to this seed case to generate the “hypos.” This includes shifting the location from school to the playground or the soccer field or shifting, “attribute[s] of the actor[s],” “the sequence of events,” and any other modifiable aspects. Rissland describes the point of this exercise as, “making sure that the set of hypos spanned a variety of possible fact situations that would emphasize and illustrate the different doctrinal approaches and sympathies of the students.” Thus, the law professor seems to intuitively understand how subtle changes to one aspect of a case can greatly change someone’s intuitive feel (or “sympathy”) of whether it counts as strong or weak.

Rissland does not go into detail on the defining process, however. My detailed analysis of the process will provide some insight into how some of these tacit influences are uncovered by students.

**Related Research: The Process of Crafting Representations**

Although there are not many general frameworks around crafting definitions in STEM education, there is some related work on the process of crafting and judging representations that involve “meta representational competences” or MRC (diSessa et al. 1991; diSessa 2004; Azevedo, 2000). One finding of this work was that middle and high school students brought incredibly rich criteria to the process. Research on MRC identified and named many of these criteria such as transparency: that a representation should need little explanation” (diSessa et al. 1991). Since crafting definitions is a related design process, we might expect criteria to play a similarly important role in defining. One study in college mathematics education gives us further evidence for this.

Although criteria were not their explicit research focus, Zandieh and Rasmussen (Zandieh and Rasmussen, 2010) touch on their importance in their work
building up their DMA framework. They present data showing that college students brought many interesting criteria into an open-ended process of defining \textit{triangle} with some striking similarities to the work in MRC. For instance, diSessa et al., in looking at how middle school students construct representations of 1-D motion, noted that students invoked the criterion of economy: “all three aspects of motion can be derived from any two that are presented” (diSessa et al., 1991). Zandieh and Rasmussen provided transcript data from a debate between two students that was surprisingly similar. The students had already included the idea that triangles have three sides and three angles. One student brought up a criterion that could be called “economy” in arguing to not include the idea that the sum of the angles in a triangle is equal to 180 degrees. Specifically, the student said: “\textit{I think the properties are like the 3 sides and the 3 angles...the angles equaling 180 seems to be derived from those properties...that's why I won't include the 180 degrees...}” Indeed, Zandieh and Rasmussen noted the important role of criteria: “The point we draw from these excerpts is that even when these students used a well-developed concept image to create a definition, their activity involved the mathematically deep issue of negotiating criteria for why certain elements (e.g., sum of the angle measures) should or should not be part of the definition. The significance of this type of activity for students is that it can contribute to developing a meta-understanding of mathematical definitions.” In this dissertation I will take a similar MRC approach of identifying and naming any criteria present in my data, adding to our understanding of how college students use criteria in the process of crafting definitions.

This review of the STEM education literature indicates that there is relatively little research on building analytic frameworks for understanding the detailed process of crafting definitions. Since the process is not well understood, it is important to first understand where students might start, unaided by frequent teacher interactions.

Related research provides some suggestions regarding what might be important in building an analytic framework. Research on MRC suggests that criteria will be important. Research on crafting law definitions suggests that an interplay between the definitional phrasing and a related set of cases will also be important. Since the law education research occurred in a classroom where the professor heavily scaffolded the conversation, it will be interesting to see how this interplay might happen unaided by a professor.

\textbf{Dissertation Overview}

In this dissertation I focus on two main research questions related to the process of crafting and using definitions:

(1) What is important to the process of definitional refinement (i.e. when words in the definition are changed or clarified)?
(2) What is the role of the definition in helping people classify new candidate examples?

To address these questions I draw on two related data sources. First, I have designed, implemented, and video-taped an open-ended workshop where undergraduate physical science majors worked to define a phenomenon with the
purpose of helping younger people identify it in the world. This workshop had minimal facilitator interaction. Additionally, I designed, implemented, and videotaped one-on-one interviews with different undergraduate physical science majors. In these interviews I also asked them to craft a definition but could follow up with them in more detail around how they might classify a set of new candidate examples.

Ultimately, this dissertation will provide a framework for understanding the process of crafting definitions and a language for talking about it. I will bring my reader through the development of this framework and its various elements. I will also:

- Name and describe a set of criteria that students bring to crafting definitions
- Name and describe a set of productive practices students engage in with respect to example use
- Differentiate the different types of knowledge that students bring to making sense of new candidate examples, which will suggest some general heuristics for definitional refinement
- Explore the educational implications of engaging students in crafting definitions
Chapter 2 – Workshop Design and Data Collection Methodology

Summary of Chapter
This chapter outlines my data collection methodology as informed by my goals for the research. I first summarize the top-level goals. I follow this by a description of what I mean by threshold and a summary of what happened during the workshops. Next, I examine many of the design variables I had to make choices around, explaining my reasoning and justifying my choices. I then provide the script that I used during the workshops as well as the materials that I gave to the students to use. Lastly, I describe the data collection methodology that involved the collection of video and audio of the workshops as well as artifacts that the students created.

Overall Goals of the Work
(1) Build an analytic framework for what is important to the process of crafting definitions, adding to the research literature in an under-studied but important area.
(2) Give some practical advice to instructors wanting to design their own curriculum around crafting definitions.

Summary of Workshops and Description of Threshold
In my workshops, participants are asked to craft a definition of the phenomenon of threshold. If time allows, they are also asked to name sub-categories and craft sub-definitions of the phenomenon as well.

What is Threshold? What contributes to people’s experience with the phenomenon?
When I use the word threshold, I don’t mean a precise physics version of the word. Rather, I’m referring to a general phenomenon that is ubiquitous in the physical and social world. Threshold-like phenomena prototypically are causal events where there is a repeated action connected to a (sometimes metaphoric) increasing amount that results in a sudden change. A human observer experiences the sudden change in a sensory way, often visually or aurally, with the change involving destruction of some kind. Destruction is not necessary but results in heightened sensory input for the observer, making types of events with its occurrence more noticeable and memorable. For instance, someone might pull a rubber band more and more, increasing the tension of the band, until it snaps. Snow might build up more and more on a roof until it collapses. Metaphorically, work might become more and more of a burden until one feels the need to collapse. I will use the term threshold-like phenomena to refer to causal events that share some or all of these prototypical aspects.

Most, if not all, humans experience many examples of threshold-like phenomena as they grow up. Kids often enjoy balancing more and more objects on top of one another until they become unstable and fall. Many languages even have idioms that speak to this universal experience. “The straw that broke the camel’s back” is found in many languages, as is “the drop that overflowed the bucket.” Thus, threshold connects to a set of rich experiential phenomena common enough to warrant an idiom in many languages.
The word *threshold*, in American English, is commonly used to refer to a (sometimes metaphorical) point in time or space. I would expect my participants to have some knowledge of the common usages of the word threshold. Thresholds are often discussed as real or metaphoric doorways. One might be on the “threshold of adulthood.” Merriam-Webster’s online dictionary currently defines *threshold* in both ways: “a piece of wood, metal, or stone that forms the bottom of a door and that you walk over as you enter a room or building” and “the point or level at which something begins or changes” (*Merriam-Webster’s online dictionary*, n.d.).

A complementary way of conceptualizing *threshold*, impacted by ones’ prior physics instruction, is to see it connected to a particular value or derivative of some variable involved in state change. State change is an idea frequently explored in undergraduate physics courses and thus college physics students have this additional resource on which to draw. Two state change examples from undergraduate physics would be the freezing point temperature when water is able to begin transforming into ice or the relative amount of two chemicals when they react to form a precipitate.

**Overview of Workshops**

**Workshop and Interview Design Task**

Workshop and interview participants attempt to meet the design goal of crafting a definition and sub-definitions of *threshold* to support younger people in identifying *thresholds* out in the world. I frame this task to the participants as an analogy to a book about birds. Such a book might include a central definition of birds having beaks and wings and feathers and being warm-blooded. The book might also include chapters corresponding to sub-categories such as large flightless birds (e.g. emus and ostriches) or colorful birds (e.g. rainforest birds like parrots). I ask that participants focus primarily on the definition in their *threshold* book and explore chapters only if there is time left.

**Developing An Example Space**

To get the participants started in this work, I first support them in developing an example space before presenting the book task. Without yet naming the word threshold, I ask them to observe and briefly describe four threshold-like examples that I provide: pushing a block off of a table until it falls, bending a pencil until it breaks, blowing a balloon until it pops, and a tea kettle being heated on the stove until it whistles. I ask if they notice any similarities between these examples.

Once they share what they notice, I use the word threshold for the first time and ask whether they feel comfortable with my use of the word threshold to describe the set of examples I’ve presented. No participants noted any discomfort with this word. I then ask the participants to brainstorm more examples of threshold-like phenomena individually. In three-person workshops, I ask groups to share out examples with one another. If students are unsure whether social examples should be part of the activity, I make a move to validate social examples. I share that a pilot study with graduate students involved the example of, “poking someone until they punch you.” This move often brings out additional social examples from participants. In one-on-one interviews, I similarly ask the
participant to brainstorm on their own and share out examples with me. I then provide the participant with additional examples from prior workshops to help fill out the example space further.

Facilitator/Interviewer Interaction

After the example space is developed, I present the design task and ask the participants to get started. Both small group and individual participants spend time crafting an initial definition and I leave them alone to do this work. The groups proceed with minimal facilitator interaction; I return only twice in the two hour workshop to ask for clarification and to ask them about a few examples that had been debated in pilot work. With the individual workshops, once the participant has crafted an initial definition, I am in continual conversation with them for the rest of the roughly hour and a half long workshop. I have a list of examples that I present to them one by one for discussion. I position the examples as ones that may or may not be a good fit for threshold. We then converse about each example and I ask follow up questions on how they are categorizing the example. Individuals sometimes choose to change their definition or name and define a new subcategory during this dialogue.

Design Goals and Design Choices

My two top-level goals for this work involve secondary design goals that guide my design choices. In this section I will list the secondary design goals associated with the top-level goals. I will then list many of the workshop design variables that I had to make choices around and my reasoning in making such choices.

Secondary Design Goals Toward Achieving Rich Data Toward Overall Goal (1), "Build an analytic framework for what is important to the process of crafting definitions to add to the research literature in an under-studied but important area":

(A) The design task should be engaging.
(B) The participants should have many conceptual resources for the phenomenon they are defining.
(C) The phenomenon should not have an easily memorizable textbook definition that students might rely too heavily on.
(D) The participants should be comfortable enough with one another and the facilitator to share ideas freely.

Secondary Design Goals Toward Achieving Overall Goal (2), "Give some practical advice to instructors wanting to design their own curriculum around crafting definitions":

(A) The activity should be closely related to the type of activity that would be feasible in a high school or university classroom that organizes its students into small groups.
(B) The phenomenon being defined should be important to as many levels of physics instruction as possible.
(C) The phenomenon being defined should engage students from as wide of a variety of background knowledge as possible.

Design Variable Choices

There are many potential places to dig into the research question of what is important to the process of crafting definitions. I will highlight the data collection design variables along which I had to make decisions, discuss the spectrum of possibilities under each design variable, and explain why I chose a particular place along the spectrum of possibilities.

Design Variable (1): Phenomenon to define
Choice: Threshold

The original idea for studying threshold came out of Andrea diSessa’s Patterns Research Group at the University of California, Berkeley, of which I am a member. This group has spent over a decade working with middle and high school students in naming, defining, and experimenting with physical patterns such as threshold, equilibration, and oscillation (diSessa and Lewis, in preparation). The phenomenon of threshold has many strengths that make it an ideal candidate for study in the process of crafting definitions.

Threshold-like phenomena weave throughout traditional high school and college physics instruction. They can be found anywhere from state change and critical points in thermodynamics to escape velocity and force balancing in mechanics. This lack of being explicitly called out by name, yet also ubiquitous in traditional instruction, makes consideration of threshold-like phenomena novel and engaging to physical science college students. In weeks following my workshops, some participants remarked, “I see threshold everywhere now!” By the end of the workshops many participants started to see connections across their undergraduate courses with one group attempting to graph and compare the torque on a block being pushed to a point where it falls off of a table and a phase diagram for water freezing. More of these details about this can be found in the first section of my conclusion chapter.

Additionally, threshold does not have a typical textbook definition that students might lean on too heavily. Work in physics education suggests that students can fall into a mode of trying to recall a memorized definition rather than sense-make around a phenomenon (Redish, 2003). With threshold, participants never made any explicit references of the need to get a “right answer”; they engaged in the definitional task according to many criteria that they believed were important for their purpose. One student said of the activity, “It was really interesting, you don’t really think about what it means for something to be something or not be something [in most physics courses].”

A final upside of threshold is that students with a wide variety of backgrounds have conceptual resources to think about it. The Patterns Group has documented middle school students productively engaging in defining threshold-like phenomena (Swanson, in preparation). This makes developmental comparison possible in future work.
**Design Variable (2):** Participants’ purpose in crafting a definition  
**Choice:** To aid younger people in identifying threshold phenomena out in the world

The purpose of crafting the definition in the workshops has a major effect on the final precision of the definition. The amount of time participants have to work on the definition also affects precision but I will discuss workshop timescale as a design variable itself later.

The definitional purpose of my workshops lies along the middle of the spectrum of more to less precision. At the most precise end might be something like working on an operational definition that aids a physicist in making an experimental measurement. At the least precise end might be an informal clarification discussion between two classmates around the meaning of a word. The middle of this spectrum is a good place to start since it likely to have some process in common with the whole spectrum and there is very little work in this area already. A visual of this spectrum is given, below.

![Spectrum Diagram]

*Figure 1.* The purpose of engaging in defining affects how precise the participants believe the definitional phrasing needs to be in a defining activity.

Defining for “younger people” meant that the participants did not engage too deeply in the mathematics underlying threshold-like phenomena. However, a small re-framing can support students in making sense of the mathematics if so desired. I explore a re-framing and some potential resulting mathematical benefits in the first section of my conclusion.

The goal of helping younger people identify an important phenomenon was an engaging task for physical science majors. Indeed, many of the participants expressed interest in leading this type of workshop themselves, either supporting freshman college students in making connections across threshold-like phenomena or putting together an online resource to support younger people in identifying threshold.

**Design Variable (3):** Facilitator Interaction  
**Choice:** Minimal in small group workshops; constant in one-on-one interviews

Minimal facilitator interaction into a group’s work during a design task allows the researcher to make claims about what students are capable of unaided
and what practices they naturally engage in. Since very little research has been performed in the area of crafting definitions, I wanted to first name important elements of what students naturally brought to the process. I therefore chose to start with two small group workshops with minimal facilitator interaction.

One of the downsides of minimal facilitator interaction, however, is that it can be difficult to make claims about the motivations for what students are doing in video data when they are not prodded to explain their thinking. Once I started to build an initial framework for the elements important to crafting definitions and developed research questions around these elements, one-on-one interviews allowed for the ability to ask follow-up questions and probe reasoning.

After I developed my analytic framework through analyzing the data, it emerged that each data set had certain affordances with respect to my analytic framework. Two central aspects of my framework are criteria around what makes a good definition and examples. The small group data, despite minimal facilitator interaction, turned out to be rich in identifiable criteria. Each of the three students in a group brought different and diverse criteria into the process such that they would often have to argue or convince one another of what criteria were important. This resonates with the research on Meta Representational Competences where many criteria around judging representations were identified during a middle school class’ whole-class discussion (diSessa et al, 1991) or small group discussions of high school students (diSessa, 2004). The one-on-one interviews were best with respect to researching students’ sense-making process in categorizing new candidate examples and the role of the crafted definition in this work. The interview format allowed the facilitator to ask different participants about the same examples so that cross-case analyses were possible. It also allowed the facilitator to better probe reasoning around the categorization process. It was much more difficult to make arguments around how students were making sense of new candidate examples in the small group data.

**Design Variable (4): Number of people working on the task**

**Choice:** Small group (3 student) workshops first, then one-on-one interviews

There are roughly three possible choices for the number of people working on the task: individuals, small groups (3-4 students), or an entire classroom of students (15-30). I chose to take a mixed approach, first taking workshop data with two small groups of three students and then taking interview data with two individuals (who had not previously been part of any small group work). This mixed approach helped with staying relevant to activities that would take place in real classrooms. The small group work closely mimicked what a large classroom environment would look like without the extra difficulties of managing so many students at once, though it did mean the loss of a whole-class discussion where students compared definitions. The facilitator interacted only twice with the students during the two-hour workshop. This would be similar to the frequency that an instructor would interact with a small group in a classroom. One-on-one work, then, allowed the researcher to follow up more deeply on the initial classroom-like format.
**Design Variable (5):** How many small group and one-on-one workshops?

**Choice:** Two of each

The choice of how much data to analyze involves breadth over depth trade-offs. In the detailed work of building an analytic framework from the ground up it is not possible to analyze hundreds or even tens of two-hour videos. Building a framework from a single group, however, is problematic also as it can be idiosyncratic and much is gained from cross-participant comparisons in naming framework elements and practices. Limiting my data to four workshops with eight total participants meant that I could research the moment-by-moment process but also have some ability to gain insight through comparison. There were a number of instances where comparison was critical in my ability to “see” and name practices. Prior work on identifying Meta Representational Competences (diSessa et al., 1991) similarly began with a limited (in the sense of number of classrooms studied) but rich data set: one classroom of middle school students over a number of class sessions in creating representations of 1-D motion. Azevedo continued to research Meta Representational Competence in a new context of map-making. He similarly drew on a small data set of interview data from three groups of three students. This supported his detailed work in naming and identifying constructive resources and criteria that students brought to the map-making endeavor (Azevedo, 2000). Video still-shots of the entire data set of two small groups and two individual workshops are shown, below:

**Figure 2.** Group A included Students M, L, and N, clockwise from left

**Figure 3.** Group B included Students A, B, and C, clockwise from left

**Figure 4.** Student G, on the right; Interviewer on the left
Design Variable (6): Student population
Choice: Undergraduate upper-division physical science majors who were associated with the same educational program.

My own background affected what populations I felt comfortable working with. I have ten years of experience instructing, developing curriculum for, and researching undergraduate and graduate students in the physical sciences. Unless there was a strong reason to study a younger population, utilizing this strength made sense. Others in my research group are currently researching middle and high school students in the process of naming, defining, and experimenting with physical patterns so comparison studies will be possible. However, whether or not comparison studies are performed, understanding the definitional practices of undergraduates is important to undergraduate education generally.

Pilot work with students along the spectrum of freshman to graduate students in the physical sciences helped me decide to narrow on upper-division students. Physics graduate students presented a difficulty in that many of them took a physics-research approach to defining, not working toward the “for younger people” part of my task, and settling on defining threshold through examining second derivatives of graphs of system variables. As mentioned in my prior discussion of the purpose of the design task, my goal was to look at a process less precise than defining toward physics experimentation. Graduate students had too much background knowledge around state change and critical points. Freshman students did not take a physics research approach to defining and also had many productive approaches they brought to the process. However, the upper-division students I worked with did not take a physics research approach and were somewhat more articulate about their process than the freshman students. The more articulate students were, the easier it was for me to build a framework. Thus, I focused on upper-division students.

Another important aspect of choosing the subject population involved needing students to be fairly comfortable with one another and with me. This was incredibly important so that students felt comfortable sharing ideas and engaging in the process. All of the students except for one\(^2\) were students who were familiar with me and with one another through our mutual involvement in an educational program at our institution. While this was not an intentional choice when I started out – students who know me are more likely to volunteer two hours of their time – I believe it to be important to the richness of the data. Azevedo, in his work on

\(^2\) This student was a friend of the other students.
identifying Meta Representational Competence in map-making, similarly highlighted this aspect of his data: “Members of each group had been classmates and seemed very comfortable with one another, as well as with the interviewer and camera operator” (Azevedo, 2000).

There are both strengths and limitations of my subject population choice. My subject population is by no means a “representative” sample of college physical science majors. In addition to their participation in an educational program together, a program that takes an explicit focus on building metacognitive skills, these students came to university from the top 10% of their high school classes and many went on to pursue physical science graduate school upon graduating. However, a representative sample is not necessary for my research questions that involve building a detailed framework for others to build upon. Having students with as many skills as possible in an ideal environment is a strength. It will be an interesting research question to determine how many students of a more representative sample of college students engage in similar practices to my limited data set and whether my analytic framework holds. My hypothesis is that a more representative sample of college physics students might engage in practices quite similar to the ones seen in my data because definitional crafting is not typically taught as part of traditional undergraduate physical science curriculum. However, my population of students does have strong metacognitive skills that certainly help them in the process of crafting definitions.

**Design Variable (7):** Amount of time participants engage in the task

**Choice:** Two-hour group workshops, One-and-a-half hour individual interviews

The choice of workshop and interview time followed from my experience doing pilot workshops. I started with a one-hour workshop and the students had many interesting ideas that they did not have enough time to fully explore. Two hours allowed for participants to work on refining their definition in more ways while not running up against a loss of engagement. Some groups showed a small loss of engagement in the last ten minutes of the workshops; others stayed late to continue to discuss ideas. This made two hours ideal for my subject population.

The amount of time participants are given affects the precision of the definition. I wanted the precision to be in the middle of the spectrum shown in Figure 1, above, so that I could analyze an extended process where a number of episodes of definitional refinement would occur. Precision also interacts with a number of other factors such as the purpose of the activity and participants’ background. This meant that it was not immediately obvious how long to make the workshops and pilot workshops were necessary to make decisions along these many interacting factors.

**Workshop and Interviewer Script**

In this section I provide the script that I followed in two parts. The first part gives the script I used for the beginning of both the workshop and interview to introduce the purpose of the activity and to help generate an example space for the participants. The second part gives the two different scripts for the subsequent
interactions with the small groups and one-on-one interviews since they are quite different.

First I will summarize the introductory script. I have the participants observe four threshold-like examples, describe them, and note any patterns they see. I then bring up the word *threshold* for the first time and say that I think all of the examples are examples of it. The participants are then asked to brainstorm more threshold examples and share them out with one another. In the one-on-one interviews, I provide some extra examples from prior participant brainstorming so that they are working with a similarly sized example space as the groups. Then, I explain the children’s book purpose of crafting the definition and sub-definitions and ask the participants to get started on crafting a definition. They do this work on their own.

For the subsequent small group participant script, I only interact with them twice during their work, asking them to clarify their work and to ask them about a few selected examples. In the one-on-one interviews, I have a cache of examples that I ask them about and we have conversations sparked by these examples. I ask follow-up questions to better understand their thinking, often choosing to change an aspect of an example in a particular way to see if participants consider it to be a better fit.

I frequently paraphrased the script instead of reading directly from it to give the activity more of an informal feel, but the paraphrasing was quite similar. For instance, the first line of the script says, “*First, I’m going to show you some videos and you’ll watch me do two things. I’d like you to write a sentence or two describing what you observe happening for each of these four cases.*” In one of my workshops what I actually said was, “*I’m going to show you four things, and for each of them you’re just going to write down a sentence or two for what you see and then I’ll explain what’s happening.*” The script was next to me at all times so I could glance down at it. The introductory script was almost exactly the same for both the group and individual workshops except at the end where the interviewer gave the participant some additional examples to fill out the example space.

**Introductory Script For Both Small Group and Individual Participants:**

The introductory script takes roughly fifteen to twenty minutes. The facilitator’s words are italicized with a description of the subsequent activity given in brackets. A copy of the worksheet that participants filled out when the facilitator prompted them to do so is provided, below:

---

3 In Group A’s workshop I asked them to note these verbally; in Group B and both one-on-one interviews, there was an explicit place on the written worksheets for them to write this down to better keep track of individual differences between noticing patterns across the examples.
Facilitator:  *First, I’m going to show you some videos and you’ll watch me do two things. I’d like you to write a sentence or two describing what you observe happening for each of these four cases.* [See Figure 6., above, to see the one page worksheet that they wrote on.]

[Facilitator shows four examples, waiting a minute between examples to allow participants to write down their sentences. Examples are: pushing a block slowly off the edge of a table until it falls, taking a pencil with two hands and bending it which eventually snaps it in half, a tea kettle being heated on a stove until it whistles, and a balloon being blown until it pops. The facilitator herself used a real block and a real pencil and pushed it over the table or broke it, respectively. The latter examples were shown as YouTube videos (Adam Willis, 2010; Moongwa, 2006) on her laptop to the participants.]
Facilitator: On your worksheet, there’s a question that asks whether you’ve noticed a pattern in this set of four examples. If you have, take a moment to write it down.

[Students write]

Facilitator: I’ve shown you four situations that I think are pretty good representatives of the idea of threshold. What I’d like you to do before we talk together is take a couple of minutes to write down, just brainstorm anything, any other examples that come to mind that you think ‘yeah, that could probably be called threshold.’ Go quickly and toss out ideas; we’ll come back together for sorting them through.

[Students write down their brainstormed examples]

Facilitator: There’s clearly many, perhaps even better, examples than what I’ve just shown you, and the idea is that we’d really map out the idea of threshold. The idea is that we’re going to write a book to help people a little younger than us identify threshold out in the world. You can imagine a book on birds has some central definition: a bird has a beak, wings, is warm-blooded, etc., and then there might be some other chapters on more atypical birds like large flightless birds (mus, ostriches) or tropical birds (particularly colorful birds like parrots). That’s the idea here – you’ll first craft a central definition and if you have time, you’ll work on the chapters. The definition can take awhile so don’t worry if that’s all that you get to. So that’s your task. I’ll play a facilitator role -- just making sure you’re staying on track, suggesting a few possibilities here or there, asking a few questions, etc. Every time I talk with people about these ideas I learn something new, so I’m looking forward to the conversation.

Facilitator*: First, I’ll ask everyone to go around and give two examples each so we have some base to work with. And then I’ll ask you to debate with each other -- are all of these, are all the ones I showed, good examples of threshold? How should we really define it?
[Students go around sharing one example out at a time. If students seem unsure whether or not social examples should be included, facilitator tells them: “A group of graduate students who did a workshop like this also came up with the examples of poking someone until they punch you.” This move either causes participants to explicitly ask, “Should we consider social examples, too?” or to start sharing social examples.]

[Facilitator hands out two pieces of paper. One has “Definition” written across the top. The other has “Chapter Ideas” written across the top. There is some scratch paper around students can write on also (both a large piece of paper the covers most of the top of the table and a few smaller 8x11 sheets)]

---

4 In the case of the one-on-one interview, the interviewer provides further examples.

Interviewer: Can you share out the examples you brainstormed? [pause to allow participant to do so] One thing I thought might be helpful before getting started: I was going to give you some examples from other people I’ve talked to, of their brainstorm. I think that sometimes they may have debated about these examples, but this is just to get more ideas out there. Other people said: pulling a rubber band until it breaks, Jenga, how high you can swing on a swing before it flips over, drinking milk until you vomit, poking someone until they punch you in the face. You can go ahead and get started on working on the definition now.
Facilitator: The idea is that as a group you’ll decide what your definition and chapter ideas are and you’ll write them on these two pieces of paper here. You’re free to use scratch paper in your work as you proceed.

[Participants begin work on crafting definition]

**Script for Interaction After Participants Craft Definition**

Following the participants crafting the definition, the script was much more free form. I had a few heuristics to follow and 2-3 particular examples that I wanted to bring up with the groups. For the one-on-one interviews, I had a list of examples that I wanted to discuss with participants but the subsequent conversation was not scripted. In this section, I’ll list the types of heuristics I used in questioning participants and the set of examples that I asked to both group and one-on-one participants.

**Small Group Script for First Interaction**

After the group worked for roughly thirty minutes, the facilitator stayed with the group for between five to ten minutes in the first interaction.

**Facilitator:** So what have you been up to?

[Open-ended discussion where facilitator might ask for some clarification around the definition, but the main goal is to have students articulate what they have been doing.]

**Facilitator:** I was going to throw out a few more examples to think about. Someone in another group had mentioned sneezing. What do you think about that one?

[Discussion. Facilitator follows up with one or more of the following examples: milk going bad and getting old.]

**Small Group Script for Second Interaction**

The second interaction was roughly another thirty minutes after the first one, but the facilitator attempted to not interrupt productive work and kept an eye out for times when work seemed to be winding down. The facilitator stayed with the group for between five to ten minutes in the second interaction.

**Facilitator:** So what have you been working on?

[Discussion. Facilitator follows up with one or more of the following questions: Has anything been especially confusing in working on the definition? Are there certain examples that seem harder to recognize somehow? What did you find most interesting in your work? What were your biggest disagreements? Would this be a good activity to do with freshman physical science students? Why or why not?]

**Examples Asked to Interview Participants (No Script)**
After the interviewee had been given roughly ten minutes to craft their definition (or more, if they wanted more time), the interviewer began to ask the participant about a set of examples. The interviewer had a list of examples on hand to draw from. This list of examples was chosen because the examples had not been immediately accepted or rejected in small group and pilot workshops and had thus led to much discussion (often with some kind of definitional refinement or naming of a new subcategory).

In both of the individual interviews, the interviewer started off the conversation with the example of milk going bad, asking each participant what she thought about it. From this point onward, however, the order in which examples were drawn upon was varied. The interviewer had a goal of asking about the following four examples to help with cross case analyses:

- Sneezing
- Getting old
- A traffic light turning and cars moving
- Hammering a nail into a piece of wood\(^5\)

However, the examples could be brought up either at strategic occasions or after a discussion had tapered off.

The full list of examples on hand also included “turning twenty-one” and “water freezing.” The interviewer, as a result of her own example generation on the fly toward strategic probing, introduced one other example. This example was an incremental version of the idiom “the straw that broke the camel’s back.” In this incremental version, the camel’s back does not suddenly break. Instead, each new piece of straw makes the camel bend incrementally until it ends up lying down on the ground. The participants themselves introduced many examples but we will not highlight them here.

In the table, below, I list the set of examples that the interviewer used in the order that they were discussed with each student. Student G and Student Q are two pseudonyms for the two participants in the interviews.

---

\(^5\) The hammering of a nail example was technically on hand for Student Q only; Student G had raised the example herself and the facilitator wanted a comparison case.
<table>
<thead>
<tr>
<th>Order of Example Introduction by Interviewer</th>
<th>Student G Interview Examples</th>
<th>Student Q Interview Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Milk Going Bad</td>
<td>Milk Going Bad</td>
</tr>
<tr>
<td>2</td>
<td>Sneezing</td>
<td>Getting Old</td>
</tr>
<tr>
<td>3</td>
<td>Getting Old</td>
<td>Hammering a Nail into a Piece of Wood</td>
</tr>
<tr>
<td>4</td>
<td>Turning 21</td>
<td>Sneeze</td>
</tr>
<tr>
<td>5</td>
<td>Traffic light turning and cars moving</td>
<td>Incremental version of “the straw that broke the camel’s back”</td>
</tr>
<tr>
<td>6</td>
<td>Water freezing</td>
<td>Traffic light turning and cars moving</td>
</tr>
<tr>
<td>7</td>
<td>Incremental version of “the straw that broke the camel’s back”</td>
<td>N/A: Not introduced by Interviewer</td>
</tr>
</tbody>
</table>

*Table 1. The order that the examples were introduced by the interviewer.*

As one can see in this discussion of the workshop and interview scripts, the introductory script was nearly identical for both cases. Subsequent interactions were more variable and took very different forms depending on the workshop or interview context. Most notably, there was constant interaction between the interviewer and the interview participants and only minimal interaction between the facilitator and the group participants.

**Artifacts Collected**

I collected the following artifacts from participants’ work during the workshop: the introductory example worksheet (Figure 6), the 8.5x11 sheet of paper with “definition” written across the top (on which students wrote their final definition), the 8.5x11 sheet of paper with “chapter ideas” written across the top (on which students wrote their final chapter ideas), and any scratch paper students wrote on during their time together.

**Room Context and Video Set-up**

The small group workshops took place in a physical science classroom on campus. This classroom was associated with the educational program that many of the students and the facilitator were involved in. It was a familiar space to most of the participants who had had many prior small group discussions there as part of their involvement in this educational program.

The classroom was set up for small group work and had a number of round tables that could easily fit 3-4 students around them. One table was chosen and the video camera was set up across the room with a wireless microphone placed on the middle of the table. The facilitator set herself up at a nearby table within hearing distance so that she could hear the group’s discussion. This supported her in asking clarifying questions during the two interactions she had during the small group workshop time. In the diagram, below, we see the camera set up, the two nearby tables, and the placement of the facilitator when she was both at her own table and interacting with the group.
Figure 7. This diagram shows the camera setup and where the facilitator was placed when the group was working alone, at a near-by table, as well as where the facilitator was placed when interacting with the group at the same table.

The interviews took place in an office in the education building on campus. Participants did not have prior experience with this space. Two chairs were placed at a desk, facing each other. A camera was set up at the opposite side of the room and a wireless microphone was placed on the middle of the desk (see diagram, below).

Figure 8. The camera and desk setup for the individual workshops.

Two hours of video data were collected for each group workshop and one and a half hours of video data were collected for each individual interview.

Transcript Convention
In the transcript that I provide, ellipses signify that I have jumped ahead and left out some of the conversation. If pauses were particularly long, I denoted them in brackets such as [pause].

Conclusion
In this chapter I have given an overview of the workshops and interviews, my top-level and secondary design goals, a discussion of many of the design variable decisions, the script that I followed during the workshops and interviews, the artifacts collected, and the room and camera set-up. This serves to give the reader a top-level summary of the workshops before we get into the details of particular
analysis episodes in the next chapter. In explaining the rationale of my design decisions I also hope this chapter aids other designers and researchers building workshops of their own.

Chapter 3 – Three Episodes Of Definitional Refinement: Identification of Criteria and Practices Associated with Meeting Them

Introduction

In this chapter I will present the development of an analytic framework and analysis methodology that was developed through an iterative process of naming framework elements, trying out the elements in data analysis, refining the framework elements, trying out the elements in data analysis and so on. After presenting the framework, I will analyze three episodes of definitional refinement. In the analysis I will focus primarily on identifying and naming criteria. I will also comment on some of the practices participants engage in when attempting to meet criteria, and the role of any examples brought up by participants.

Analysis Methods: Development of Analytic Focus and Creation of Analytic Framework for Definitional Refinement

Development of Initial Framework

Given the exploratory nature of this study, I first asked the question, “what is important to the process of crafting definitions?” I decided to start with identifying definitional refinement, i.e. moments in the data when words or phrases in the articulated (either verbally or written down) definition are changed. After finding episodes of definitional refinement, I narrowed my question to, “what is important to the process of definitional refinement?”

My first level of analysis of the episodes of definitional refinement involved a narrative summary (Barron and Engle, 2007) of everything leading up to the definitional refinement in which I tried to tell a story about what happened. Additionally, I used a top-down meets bottom-up approach in deciding to focus on identifying example and criteria use during the definitional refinement work. Prior research (diSessa et al., 1991; diSessa, 2004; Star and Bowker, 2003) suggested that criteria were central to design processes and worth paying attention to. I went into the data with the hypothesis that they would be important. Empirically, in my initial viewing and content logging of the data (Jordan and Henderson, 1995), I also noticed that examples were core to the process as well with one group brainstorming over 50 unique examples during their time together.

In analyzing the data, I flagged any example that was discussed. I also flagged any statements where the students explicitly discussed the word definition, a good proxy for criteria. For instance, one student made the statement “our definition has to be overarching, it has to cover everything.” Additionally, some statements flagged
as criteria did not explicitly involve the word definition but it was clear that the student was referring to it such as “is that too general?” or “that word is too complicated.” I also flagged anything that appeared to be a meta-comment around what type of activity the group should be working on.

After performing this first level of analysis, I determined that refinement was built upon two important components: commitment to criteria and skillful interaction with examples (see Figure 9, below). This arose from looking at my flagged criteria and examples and narrative summaries. I noticed that refinement happened not only because students articulated criteria but also because they were committed to them: they had to convince their group members to engage in them and then put sustained work into meeting them. Additionally, students’ invoking of examples was more than just random; they would make arguments with examples. Lacking a more precise term at this juncture, I labeled this simply “skillful interaction with examples.”

![Figure 9. Preliminary Analytic Framework](image)

**Development of an Improved Analytic Framework**

With the above framework as a base, I began a deeper level of analysis. This process had resonances with the methodology of analytic induction, which was first formulated by Znaniecki as a methodology in sociology in 1934 (Znaniecki, 1934). I deeply examined a small number of episodes and built an initial framework. This framework then allowed me to look for patterns and formulate new questions, which allowed for further refinement of the framework. For instance, my research questions were able to evolve to, “how does definitional refinement begin?” and, “how do students decide when to keep working on a particular activity involved in refinement and when to end it?”

Examining the beginning of definitional refinement more closely led me to make an additional iteration on my analytic framework. Definitional refinement often began with the crafted definition being called into question by either an example or a criterion that it did not meet. This picture called for turning my preliminary framework on its head with the crafted definition itself playing a role. The refinement would proceed with an interplay between examples and the crafted definition, mediated by criteria. For instance, an example might call the crafted definition into question, but the group always invoked criteria in deciding what to do about it and how to proceed. A second attempt at a framework is found, below:
Adding A New Framework Element

I then tried the above framework out with the episodes of definitional refinement in my data, but determined that it still failed to capture some important things. First, I noticed that there were times that my participants used knowledge outside of their crafted definition. They would sometimes come across an example that felt “off” in some way and judge it through ideas not explicitly included in their definition. If a participant could then realize that they were using ideas not articulated in the definition, this became an important way that definitional refinement could begin. It was thus necessary to add an additional element to my framework: definitionally unarticulated knowledge or DUK for short. There were now four elements in my framework: examples, the crafted definition, DUK, and criteria. In the figure, below, “criteria” was given a different color to call attention to a difference between knowledge and meta-knowledge. Examples, the crafted definition, and DUK were all types of knowledge that participants brought to the process of defining whereas criteria were used to make decisions around what to do with this knowledge.

![Diagram of framework elements]

**Figure 11. Framework elements**
Final Framework For The Process of Definitional Refinement

While adding a new framework element, I also noticed that definitional refinement proceeded in two somewhat distinct phases. There was a pre-refinement phase where participants had some debate on whether to engage in changing the definition. Empirically, this started in one of three ways. Participants would notice a misalignment between an example and the definition, a criterion and the definition, or an example and DUK. Once a group or individual decided to change the definition according to one or more specific criteria, their work involved further practices associated with additional criteria, examples, and attempts at new words and phrases. The complexity of this process was not captured by my second attempt at a framework and I made a final iteration, below.

![Definitional Refinement Diagram](image)

Figure 12. Final Framework for Definitional Refinement. D1 represents an initial definition and D2 represents a refined one. Framework elements such as examples, DUK, the crafted definition, and criteria, are part of each of the blue and green boxes are sometimes hidden within the term “practices.”

Since criteria were central to this process, I took this new framework as an opportunity to zoom in on criteria and ask, “how do I better identify criteria?”

---

6 It seems plausible that a misalignment between DUK and the crafted definition could result in the potential for definitional refinement. However, in my data, it was participants' noticing that an example did not fit with some DUK that made participants aware of particular DUK. Its typically tacit nature probably prevents the DUK/crafted definition misalignment from being a frequent contributor to definitional refinement.
two phases also suggested an additional question: “can I characterize and find patterns in the practices associated with the pre-phase of determining how to proceed and the refinement-phase of practices associated with meeting one or more criteria?”

**Better Criteria Identification**

My initial methodology of identifying criteria had been through explicit references to the definition or meta-statements. I refined and expanded this methodology by including analysis of participants’ practices, sometimes also looking for a history or future of these practices in the video data. Additionally, if a participant made an explicit criterion statement, I also argued around whether or not it was reasonable to take the participant’s words at face value.

One particular participant practice that could support better criteria identification centered on examples: changes to the definition were often checked with one or more examples. I asked the question, “why did the student bring up this particular example?” and looked at prior and subsequent episodes of definitional refinement to see if I could identify patterns in how participants were choosing to bring up examples. Sometimes participants seemed to check the definition with whatever example they could most easily recall or it was unclear how they were choosing certain examples. However, some participants repeatedly brought a small set of examples to bear on checking the definition. Although a student might not explicitly say, “The phenomenon our definition describes has a set of examples that I believe to be most prototypical and the definition should always cover these examples,” I could argue that, by her or his actions, a participant thought this was an important criterion to bring on the work of refining.

In addition to looking at practices associated with examples, practices associated with participants’ negotiation of explicit criteria could sometimes support additional criteria identification. For instance, sometimes a participant might bring up a criterion, but other members of the participant’s group would not immediately take it up. This could mean that the other group members simply did not agree with the importance of the criterion, or, it might mean that they thought a different criterion was more important. When they were not explicit, looking at their statements and practices sometimes allowed for plausibility claims around the existence of additional criteria.

**Identification of ‘Practices Associated with Meeting Criteria’**

When a group or individual decided to work on meeting a criterion, I labeled this work “practices associated with meeting criteria.” At first, I wrote up a narrative summary of the work. I then compared narrative summaries across different episodes to see if there were similarities or differences. This allowed me to better name practices if I saw them more than once.

**Data and Analysis**

**Overview**

In this chapter I will present data and analysis from three episodes of definitional refinement. The three episodes included in this chapter were selected because they were the clearest with respect to supporting my ability to identify
Definitional refinement, criteria, and the practices associated with meeting criteria. Primarily, the analyses will focus on identifying criteria. I will also comment on some of the practices participants engage in when attempting to meet criteria, and the role of any examples brought up by participants. Each of the three episodes will proceed in the following way: first, a summary of the data, next, the transcript, and lastly, the analysis. In some cases the summary and the transcript may be somewhat intertwined.

**Definitional Refinement: Episode One**

**Introduction**

The first episode involves three students from Group B, Students A, B, and C, as shown in the photo, below. All students are upper-division undergraduate physical science majors. The summary, transcript, and analysis will be split into two parts: the pre-phase (see Figure 3, reproduced, below) which begins when a participant notices a mismatch between an example and the crafted definition, and the refinement-phase, where the group works on meeting one particular criterion. To make analytic arguments about criteria in the pre-phase, some excerpts from the history of the workshop will be presented first.

![Figure 3. Group B included Students A, B, and C, clockwise from left](image)

**Excerpts from Workshop History**

The episode of definitional refinement that we are about to examine begins when one participant, Student A, notices a conflict between an example and the crafted definition. To understand how Student A came to notice this conflict, it is necessary to go back to two excerpts from the history of the group's workshop. I will show that Student A frequently engaged in a practice of checking initial ideas for or changes to the definition with a set of a few particular examples. These examples were not border-line cases; they were examples whose fit was never challenged by any group members. I will label this heuristic move **Core-Checking**. Underlying this heuristic is a criterion: that the definition should always cover examples deemed (implicitly or explicitly) to be most core or prototypical. I will label criteria as phrases that complete the following sentence: “The definition should...” In this case, the criterion would be **Encompass Prototypical/Core Examples**. This criterion, while a somewhat obvious starting point for definitions, guides participants’ work.

Student A core-checks for the first time when the group is considering their first try at the definition (“an action is applied to an object until a reaction occurs”).
Student A makes a move to compare this definition to three of the four examples that the facilitator originally showed (water boiling in a kettle, an object pushed over the edge of a table until it falls, and bending a pencil until it snaps). Student A privileges these examples in his comparison in that he does not discuss any others brainstormed by the group. This could be influenced by the fact that the facilitator affirmed them as examples. We see Student A discuss a very rough check of whether the definition makes sense with these examples and he does not express any discomfort with the fit at this time. Specifically, Student A says:

“So if we think about the examples we were shown, there was like the pencil, applying like a force, and there [for the block pushed over the table], it was a gravitational potential, being overcome, ‘cuz once it was on the edge of edge of the table, it overcame gravitational potential energy until it hit the floor?...water boiling...”

Of the three examples, he discusses the block more in-depth than the other two. Roughly five minutes later, we see Student B also engaging in what appears to be core-checking. The group is discussing how to include the idea that objects have a property that is connected to the threshold. Student B says:

“We still need to say we’re only looking at one property of the object, not just the object. So say for the balloon, we’re looking at elasticity, we’re looking at how much the plastic can hold, how much it can stretch before it breaks, so then, for each object, examine, like, one property, and push that property to the limit...”

Student B, in examining this new idea to include in the definition, tries it out with one of the four examples from the facilitator again: the balloon. A few minutes later, Student A again core checks in explaining Student B’s thinking to the facilitator. The facilitator interrupts the group and asks what they had been working on, interacting with them for about five minutes. Student A uses Student B’s balloon example again, but adds the block example as well:

“[Student B] brought up the point that in each of the examples, they had one property whose threshold was reached and gone over...say in the balloon, its elasticity or something, or in the rectangular block, the gravitational potential that was being overcome...[is] one thing we want to look at and explain.”

Note that Student A seems to be talking incorrectly about the physics of the block scenario though it is hard to know exactly what he means by the word potential. First, he seems to be talking about potential (“the potential was overcome”) as one would talk about forces. Second, if he really meant force rather than potential, he has still named gravity as the force to overcome. At some point there will be a net torque on the block causing it to fall off of the table. This occurs once rxF from the gravitational force about the pivot point overcomes the normal force from the table (rather than the other way around as Student A seems to be discussing it). Our research in this chapter is focused on criteria and practices associated with meeting them. Additionally, this is a small excerpt from the entire process so we will not pass judgment on it in this moment.
This is the second time that Student A has used the block example as central in his core-checking.

**Definitional Refinement Pre-Phase: Summary and Transcript Excerpts**

I now return to the episode of definitional refinement, which begins about ten minutes after the facilitator interaction described above. After the facilitator left, the group had been discussing the idea that objects have a property that affects the threshold. As part of this discussion, Student A makes another attempt to write down a definition. In my framework diagram, this would be considered the initial definition, D1:

"What changes is the object...each object has a property...that sets a limit on how much it can be pushed."

Next, Student A notices that the definition does not cover all examples. He says, "even the object doesn't necessarily change." Student B questions him: "really?" Student A then defends his idea, perhaps unsurprisingly at this point, with the example of pushing an object slowly over the edge of a table until it falls. He physically pushes a pen over the edge of the table and says, "I dropped this [the pen]; the object is still the same." Student B immediately gets this demonstration and agrees: "yeah, that's true."

When Student A first notices that the object doesn't necessarily change, I argue that, given previous data, it is likely that he made a quick core-checking move with the block example as he's done twice before. This would explain why he is so quickly able to draw on this example. Student B himself also appeared to core-check once which might explain why he is so easily convinced by Student A.

Next, there is some debate that occurs. Student C is does not appear as easily convinced and questions, "What about the rubber band?" The rubber band is one of Student C's own examples that she brainstormed herself when the facilitator asked everyone to brainstorm examples. Student B acknowledges her, but backs up the idea that they should change the definition:

"We thought that was always the case [that the object changed], but he just made a good point. He just dropped the pencil, the pencil didn’t change, but something about the pencil changed. Position or potential energy."

The students then continue the debate. Student C agrees with their point about the object not necessarily changing but then asks, "So then there should just be exceptions." Student B counters: "Exceptions? Or we just need to change the definition." Student A makes a move to agree with Student B saying that: "Our definition has to be overarching; it has to cover everything."

**Analysis**

I have presented evidence for the criterion of **Encompass Prototypical/Core Examples** along with its associated practice of core-checking. This particular practice may help with students’ working memory limitations and pragmatic time constraints. Checking every new definition with tens of examples
can become time consuming and, except in special cases, probably does not help much more than checking with a few examples. This practice also helps the group realize that one of their core examples does not fit under their crafted definition.

There is some evidence for two additional criteria. Students A and B state simply that definitions should cover the entire set of known examples of a phenomenon, a Cover All Known Examples criterion which covers a wider set of examples than Core Fit. Student C does not express the same idea that every example needs to be covered within the definition. Instead, she introduces the idea of exceptions. Although there is not enough data to say for sure, Student C could be operating from a Disregard Outlier Cases standpoint that would be quite rational. If one considers a theoretical situation where the group discusses 1,000 examples that they deem to fit with only one case of the object remaining the same, her position might have been taken more seriously. Since she does not argue strongly for her position, it may be the case that she does not have a strong sense of the importance that definitions should Disregard Outlier Cases and is instead simply without the Cover All Known Examples criterion and thus sees no problems with exceptions which would allow the group to move on and not engage in working on changing the definition further.

Another interesting question, albeit one that the evidence is too sparse to answer is: “why didn’t the group consider putting the block example outside of their boundary of threshold since it didn’t fit their current definition?” One answer could be that since the block example was part of Student A’s core, indeed central to his core-checking, he was less likely to question whether or not it even belonged in the category. D1 was also a fairly new attempt at a definition so may have not carried as much weight to outlaw examples. Additionally, there was never any work done to brainstorm other examples where the object did not change to see if the block example might be a fluke. Students A and B took it to be a good example without question. This all suggests that the block example is a really compelling example of threshold. diSessa said of the block example: “Is it ‘objectively’ a good example? I think it really might be such. The parameter is clear. The mechanism is clear. The example is simple even for kids” (A. diSessa, personal communication, November 16, 2013). In the next chapter we will see many examples that participants do not so readily accept when they come into conflict with the crafted definition.

Definitional Refinement Refinement-Phase: Summary and Transcript

Ultimately, Student C went along with Students A and B and the group decided to attempt to meet the Cover All Known Examples criterion. This will involve changing the definition to cover the block example. Recall that the last state of their definition was “What changes is the object...each object has a property... that sets a limit on how much it can be pushed.”

To work on meeting this criterion, first Student B attempts to see if he can just change one word: “So I guess we need to say system instead of object.” As the students attempt to make this change they run into an issue: they decide that the word system is difficult to define. Student A makes an attempt to address this by advocating for not talking about change in the definition. Students B and C counter that change is too core to their understanding of the phenomenon to leave it out. Student A tries again with a more vague phrase that includes the idea of a change
but not the word system. The group works on this phrase until they are satisfied with it. Finally, a new definition, D2, is written down and verbally articulated by Student C: “An object is pushed to a limit until the limit is overcome and something about the object changes.” The definition has changed from “the object changes” to “something about the object changes.” Their full dialogue follows:

Student B: So I guess we need to say system instead of object
Student C: The object still needs to stay there
Student B: An object is pushed to a limit at some point
Student A: until the limit is overcome
Student B: the limit is overcome and the system
Student C: and/or object changes
Student A: we have to, what is the system? If for now we stick with object, that’s fine. You don’t have to say, you don’t have to talk about there’s been changes, not necessarily, just say oh it’s like
Student B: I think you have to say change, well, ’cuz that’s the whole threshold thing
Student C: You know how [the facilitator] was saying ‘a bird has wings, a bird has a beak’ – how would you, how would someone know that threshold applies if there is no end result? [pause] I think that’s what makes it easier to identify.
Student A: Can we just be like ‘until a change happens?’
Student B: Or something about the object changes
Student C: Object or the system. Can’t say both object and system?
Student B: Yeah, but we should try to avoid system, ’cuz then, that’s hard to define.
Student A: Yeah
Student B: just say ’until something about the object changes.’ And something can be, like, for the pencil [that you push over the table], the potential energy, kinetic energy
Student C: [writing and reading the writing aloud] An object is pushed to a limit until the limit is overcome and something about the object changes.

Analysis

As the group is working on satisfying the Cover All Known Examples criterion, we see practices associated with changing the definition. The group sees if they can change just one word from the current form of the crafted definition. They also invoke other criteria as they proceed and attempt to modify words and phrases. I will identify new criteria in the order they appear in the above dialogue. Student B makes a move to simply replace the word object with the word system. However, Student A struggles with this word (“what is the system?”) and advocates not using it. One could describe Student A’s move as using a Use The Simplest Words Possible criterion though he is not explicit that this is his intention. It is somewhat implied by his phrase, “If for now we stick with object, that’s fine,” which is a return to simpler phrasing. Later, Student B will explicitly note that system is “hard to define” and Student A will quickly agree with him. This evidence strongly suggests that Student A is invoking a kind of Use The Simplest Words Possible criterion.

Student A’s next move, to not use this word system by getting rid of the idea of change, is met with strong resistance by his group mates. Student B employs a
criterion one might call **Cover the Conceptual Core**. Although he is not explicit in saying something like, “change is core to my conceptualization of threshold, we can’t leave it out!” he states that change is “the whole threshold thing,” which one could construe as an appeal to **Cover the Conceptual Core**. Student C also advocates for including the idea of change, but uses a more practical criterion of **Meet The Design Purpose**. The facilitator had initially framed the activity through an analogy of a bird book: that the definition should help people identify threshold in the world.

Student C invokes this design purpose when she states, “You know how [the facilitator] was saying ‘a bird has wings, a bird has a beak’ – how would you, how would someone know that threshold applies if there is no end result?” Without an “end result” or change, Student C does not see it meeting this goal. Student A does not further advocate for getting rid of the idea of change.

Finally, there is some debate between Students B and C around whether to use the word system, with Student B also employing the **Use The Simplest Words Possible** criterion. We see this explicitly in his phrase, “Yeah, but we should try to avoid system, ‘cuz then, that’s hard to define.” The group finally settles on the more vague phrase “something about the object changes” that balances the **Use The Simplest Words Possible, Cover the Conceptual Core**, and **Meet The Design Purpose** criteria in ultimately meeting **Cover All Known Examples**

**Conclusion**

In this episode we see how the practice of core-checking helped the group realize that one of their threshold examples did not fit under their crafted definition. There was then some debate on how to proceed: should the definition be changed or should there be exceptions? The criterion of **Covers All Known Examples** won out and the group proceeded to change their crafted definition to include this example. We saw some practices associated with meeting this criterion along with additional criteria that helped to guide the changes that the group made. I made a strong case for the existence of the criterion **Encompass Prototypical/Core Examples, Covers All Known Examples, Use The Simplest Words Possible, Cover the Conceptual Core**, and **Meet The Design Purpose**. I made a weak case for the possible existence of **Disregard Outlier Cases**.

**Definitional Refinement: Episode Two**

**Introduction**

Episode Two is drawn from data from Group A, a different group of three students. Episode Two follows a similar trajectory to Episode One: one participant in the group advocates for changing the definition according to a particular criterion but there is first some debate about doing so. The debate is resolved and the group employs some practices associated with meeting the criteria until it is sufficiently met. I will split up the data into two parts (pre-phase and refinement-phase) with analysis following each section. In part one (the pre-phase) I will show how a participant advocates for a particular criterion and the debate that follows. In part two (the refinement-phase), the group will decide to work on this one criterion and I will show the practices associated with meeting it.
Definitional Refinement Pre-Phase: Summary and Transcript

The episode involves three students, M, L, and N, seen in Figure 2, reproduced, below. It is roughly five minutes in length and comes from the first few minutes of work that the group does together on the definition after the facilitator leaves. After an initial definition is proposed, one student makes a move to have the group engage in meeting the criterion of **Sufficiently Precise As To Not Be Mistaken For Another Word**, i.e. that the definition should not be easily mistaken for another word. She says that the current form of the definition could be mistaken for *force*. Convincing the group to work toward this criterion takes some effort. This includes pushing back on one of her group members who engages in some possible core-checking. Eventually she does convince her group and they refine the definition, better disambiguating between *forces* and *thresholds*.

![Figure 2. Group A included Students M, L, and N, clockwise from left](image)

The initial discussion is directed by Student N who asks, “**What were your first thoughts? When she [the facilitator] was, like, ‘central definition’?**” Student M discusses her initial idea: “**Is it when it starts doing something else than what it’s previously doing?**” Student N agrees with her, but Student L questions: “**Is that a little too broad?**” Student M clarifies her definition: “**Yeah, like, uh, something doesn’t happen and then it happens, right?**” Student L advocates for further refinement: “**But that’s, like, the definition of force, so what’s the difference between force and threshold?**” Student M does not immediately engage with Student L, instead using two of the four facilitator examples to check their definition with: “**Oh for, like, the boiling water, too, right, it’s like not whistling and then it whistles and stuff like that, or the balloon, it’s like a balloon and then it pops.**” Student L continues to advocate for her position: “**...if someone told me the definition for threshold and the definition for force, I’d be, like, what’s the difference?**” After making this statement, her group members decide to join in with her in the task of disambiguation. The full transcript of this data follows:

Student N: **What were your first thoughts? When she was like ‘central definition’**
Student M: **Is it when it starts doing something else than what it’s previously doing?**
Student N: That’s exactly what I thought. I was like, when there’s a change of state basically. When you push something until the action is, like, changed, or something happens
Student N: Something happens
Student L: Is that a little too broad?
Student N: It might be.
Student L: um, so
Student M: Like if you’re in a kids book I guess you’d want
Student N: well, but I was just thinking in terms of where we can go
Student M: Yeah, like, uh, something doesn’t happen and then it happens, right?
Student L: But that’s, like, the definition of force, so what’s the difference between force and threshold?
Student M: Oh for like the boiling water, too, right, it’s like not whistling and then it whistles and stuff like that, or the balloon, it’s like a balloon and then it pops
Student N: And then it pops
Student L: But until I was like 5 or 6, that would be the same exact definition, if someone told me the definition for threshold and the definition for force, I’d be like, what’s the difference? I guess we can give specific examples of, like, I’m not sure how to explain the difference between force and threshold.

Analysis
In this episode we see the criterion of Sufficiently Precise As To Not Be Mistaken For Another Word employed when Student L expresses her discomfort that a child might think that their current definition could be mistaken for force instead of threshold. Instead of joining Student L in working toward this criterion, Student M engages in what appears to be core-checking of the current definition.

I can’t make as strong of a case for core-checking in this data as in Episode One since we do not see Student M engaging in this practice repeatedly in the data; it might simply be that she is engaging in checking the definition against the first examples that she could most easily recollect. Either way, she seems to be making a move counter to Student L’s suggestion, instead making a point that the definition covers some of the main examples the group has discussed, possibly making it good enough. It is not clear whether Student L is trying to push back on Student M’s criterion because she believes the definition might be good enough or whether, because it is very early on in their work, she wants to work on solidifying the group’s attempt at a crafted definition before critiquing it. It takes Student L further arguing for Sufficiently Precise As To Not Be Mistaken For Another Word again before Students M and N decide to try to work toward meeting this criterion, though Student L herself mentions that it might not be easy: “I’m not sure how to explain the difference between force and threshold.”

Definitional Refinement Refinement-Phase: Summary and Transcript
The group then follows Student L in trying to change the definition to better disambiguate force and threshold. They work on articulating the ontology of threshold, pulling apart that forces are causal mechanisms and thresholds are boundary points, a “when” in time. Student M first advocates for getting clear on the ontology of the word: “What is a threshold, right? Is it like, a thing, is it a point in time,
is it like, when two conditions are equal...” Student N proposes that a threshold is a moment when something transforms into something else, but isn’t sure that transforms is clear enough language. She puts forth an example of a person that pushes a block to transform it from stationary to moving, but she does not want to qualify this example as a threshold unless a more major transformation happens: the block falling off a cliff or melting. Student L argues back that the transition from stationary to moving is enough to be considered a threshold, gaining buy-in from her group members. They continue working on the ontology of threshold, eventually refining their definition from their initial verbal articulation, D1, of “something doesn’t happen and then it happens” and “when it starts doing something else than what it’s previously doing” to D2, “A threshold is a boundary. And forces are what lead you to cross the threshold.” The full transcript follows:

Student L: I guess we can give specific examples of, like, I’m not sure how to explain the difference between force and threshold.
Student M: What is a threshold, right? Is it like, a thing, is it a point in time, is it like, when two conditions are equal, like
Student N: More of the latter, more of what you just said, like, when it transforms into something else, I guess, but that’s still, like, sort of, like, if you push something, it transforms from a stationary block to a moving block, but it’s not a threshold until it goes over a cliff, or it, like, you push it too fast and it melts
Student L: It’s also a threshold, the fact that it goes from not moving to moving, is threshold
Student N: Oh that’s true, that’s true
Student M: Ohhhh
Student N: A force is something that makes you, approach a threshold?
Student M: So maybe a threshold is like a barrier? I dunno----
Student L: Yeah
Student N: And the forces, sort of like, the chemical thing, catalysts and enzymes and stuff
Student M: Yeah
Student N: So, force sort of leads to [short pause] threshold?
Student M: Threshold crossing, perhaps
Student N: Oh, yeah. The force leads to the threshold crossing, but the threshold is an independent thing
Student M: Yeah, I guess maybe it’s like important to have a concrete idea of what a threshold is, like, can we identify it as, I mean, like obviously it is a noun, but is it like, a physical thing, yanno what I’m saying? or is it like, like, how to, how to define it, I dunno
Student N: It’s a boundary
Student L: Yeah
Student M: A boundary, okay
Student N: That really ties into the crossing thing
Student M: Yeah.
Student N: So a threshold is a boundary. And forces are what lead you to cross the threshold. And then we should say what happens when you cross the threshold.
The group does not work any further at this point on what happens when a threshold is crossed, instead deciding to move on to a different activity of brainstorming examples for their book. About twenty minutes later when the facilitator asks the group what they have been working on, Student N recaps their work by saying, "We had a discussion in the beginning about the difference [of threshold] from force because they both sort of cause things to change, but we decided that the force is the thing that makes you cross the threshold, but the threshold is the boundary between two different states so it's like the place where things change."

**Analysis**

In their attempts at meeting **Sufficiently Precise As To Not Be Mistaken For Another Word**, there are two connected practices that are productive for the group: using both words in the same sentence and explicitly trying to identify the ontology of threshold. They are connected in that using the two words in the same sentence can help the students better identify what ontology they have been implicitly using as to critique it. Student N frequently attempts to use both *force* and *threshold* in a sentence together such as, "A force is something that makes you, approach a threshold?" and "So a threshold is a boundary. And forces are what lead you to cross the threshold." This is a productive practice that gets at the interrelationship between the two words and the group gains new insights. The second practice is explicitly working on the ontology of threshold. Student M questions, "What is a threshold, right? Is it like, a thing, is it a point in time, is it like, when two conditions are equal" She also asks, "So maybe a threshold is like a barrier? I dunno--" These two practices, in concert with one another, lead the group to refine their definition considerably in a short amount of time.

Another practice we see is a student trying out new definitional language with an example. Student N suggests a new phrase for the definition: "A force is something that makes you, approach a threshold?" Then, a few seconds later, she checks it with a particular example of chemical reactions and precipitation, an example that one of her group members had shared during the initial brainstorming session: "And the forces, sort of like, the chemical thing, catalysts and enzymes and stuff." There isn’t enough evidence to say exactly why she employs this example. It might be the one she most easily remembers. It might be that the group has been discussing ‘pushing a block’ type examples where the force is easily identifiable and she wants to consider a broader range of examples to make sure the phrasing makes sense across a broader range. Either way, the group does not see any problems with the fit of the chemical precipitate example with the new phrasing and they continue on with their work.

**Conclusion**

In Episode Two we have made a strong case for the existence of the criterion that the definition should be **Sufficiently Precise As To Not Be Mistaken For Another Word**. Definitional refinement occurred when the group engaged in meeting this criterion. We saw practices associated with this work: explicit work on word ontology as well as using two words that need to be disambiguated in the same sentence. Finally, we also saw more evidence of students consistently employing examples in their work. Students checked new words and phrases in the
crafted definition with examples: the ones most easily remembered, a more intentional core, or a diverse range.

Definitional Refinement: Episode Three

Introduction

In this episode I examine definitional refinement that occurred in one of my one-on-one interviews. The episode is considerably more directed than the previous ones since the facilitator is having a conversation with the participant rather than letting a group work on their own. The interviewer introduces a new candidate example to consider and, at times, encourages the participant to write down ideas. However, the participant drives the refinement. She voices criteria and we will also see a completely new kind of definitional refinement: deciding that the crafted definition should be broken down into two different words: threshold and breaking point.

This episode involves Student Q, an upper-division undergraduate physical science major. She is shown in Figure 5 which is reproduced, below. The episode occurs after the student has written down her crafted definition and the interviewer has asked her about three candidate examples of threshold that had been debated heavily by previous participants. The conversation about these three examples takes about ten minutes. Then, the interviewer asks about a fourth new candidate example, sneezing, and this touches off the discussion and definitional refinement I will examine here.

Figure 5. Student Q, on the right; Interviewer on the left

This episode is slightly more difficult to break into the pre-phase and refinement-phase due to the interaction with the interviewer. One might call almost the entire episode the pre-phase since it is not until the very end of interaction that Student Q invokes an Consider Its Audience criterion in deciding that it is necessary to include both threshold and breaking point into the definition when she says: "If you introduce it [threshold] you need to acknowledge both [threshold and breaking point] because the person reading it would be like "what about the other one?" However, due to the interviewer’s suggestion that she write ideas down along the way as they made sense of example together, Student Q had already written down an additional phrase into her definition to highlight this distinction. Thus, the
pre-phase and refinement-phase are more mixed in this episode and I will present the whole episode without breaking it up into the two phases.

**Summary and Transcript**

I will first summarize and provide transcript excerpts from the initial conversation the interviewer and Student Q had about sneezing. Then, I will provide a summary and more detailed transcript data from the parts where Student Q lays out particular criteria and works on refining.

At the point in the interview when the interviewer first asks her about the example of sneezing, Student Q had the following written definition for threshold, D1: "A finite point where something turns into something else, or the point at which something happens when you act on an object." She includes the idea of a state change as well as an action that results in reaction. When asked about sneezing she does not immediately accept it as a good example. She explains that her sneezes surprise her because she does not experience a build up. The idea of build up is not found explicitly in Student Q’s definition, but build up becomes an important factor in making sense of examples for Student Q. She compares sneezing to coughing, with coughing as a better example due to the existence of a clearer build up:

*There's a point at which you have to sneeze and, let's say it's more like coughing, you can take a little bit of annoyance in your throat to the point where you're like coughing up organs and stuff and there's kind of like a trigger point almost, whereas, I guess there is, too, for sneezing, but, I dunno, my sneezes always catch me by surprise usually...There's no build up, almost, well, probably there is, but not that I've ever noticed.*

The interviewer and Student Q then proceed to have a conversation about the various causes that might be involved in sneezing for about a minute. Student Q acknowledges that a cause like pollen could be built up, but that there are many other possible factors. The interviewer attempts to close the conversation by laying out this confusion: "Anyway, all that to say, sometimes, if there were really a build up of something, it would make it in [to the book], but given that we don't totally know how it works..." Instead, Student Q brings in another example she had previously brainstormed during the initial brainstorming phase, walking through a door, to question the very necessity of a build up. She eventually articulates that there are two ways to think about threshold, one involving a build up to a barrier and another of just the barrier itself. Specifically, she says:

Student Q: *But at the same time, if you're talking about a threshold being a door, is there a build-up to the door? I guess you're walking toward the door--*
Interviewer: *Yeah, that's a good question*
Student Q: *There's not signs, like 'there's a door here, get prepared!' -- you're about to have a door! So I dunno--*
Interviewer: *Yeah, that's a good point, so it's like, there's something, in terms of the way we were talking about it, it seems the walking to the door is an important part of it*

40
Student Q: Yeah, I guess [upward intonation expressing doubt]
Interviewer: Or maybe it's not? I dunno
Student Q: I dunno, it just seems like, it's kinda hard to define, because it's one of those words that has a little bit of wiggle room, and they're kinda more abstract concepts where you can't define it completely
Interviewer: Yeah, like there's people who, say, talk about it in different ways?
Student Q: Yeah. It's kinda one of those things that not everyone is going to 100% agree, because everyone is going to have slightly different interpretations of what threshold is. I dunno, unless you're someone who's like crazy about categories, everyone's pretty much gonna agree, that's an apple, that is not. I dunno, though, tomatoes seem to be confusing.
Interviewer: Whether they're a fruit or vegetable?
Student Q: [nods]
Interviewer: Yeah, so you're kind of saying people could define threshold with something leading up to the door or just the door
Student Q: Or just the point where, the sneeze and then not sneeze, the sneeze
Interviewer: Maybe that's worth writing down, too, so you're saying there's these two ways of seeing it...

Student Q takes a moment to write down her idea. Specifically, she writes: “2 ways of seeing threshold: some point that is built up to, or just a barrier. Snapping a pencil has a build up of strain. A sneeze is just not not sneezing.” She then makes a move to attach two different names to each idea, threshold and breaking point, and considers her audience in deciding that including both ideas is important. Specifically she says:

Student Q: I guess you could call maybe one a threshold and the other a breaking point almost?
Interviewer: Yeah, sure, I like that distinction
Student Q: If you introduce it you need to acknowledge both because the person reading it would be like "what about the other one?"

The conversation about sneezing ends at this point and the interviewer moves on to asking her about a new candidate example.

Analysis
In this data we see how the discussion about one example, sneezing, that appeared to lack a clear build up, eventually resulted in Student Q refining her definition. In the prior two episodes of definitional refinement in this chapter, it was a misalignment between the crafted definition and either an example or a criterion that made the refinement possible. Here, however, it is a misalignment between an example and definitionally unarticulated knowledge or DUK that Student Q articulates. She uses the phrase “build up,” not found in her definition, in articulating what feels “off” about sneezing. She then refines her definition to include this idea. Specifically, she refines her definition from D1: “A finite point where something turns into something else, or the point at which something happens when you act on an object,” to D2 which is the same as D1 but with the additional
phrase, "2 ways of seeing threshold: some point that is built up to, or just a barrier. Snapping a pencil has a build up of strain. A sneeze is just not not sneezing."

I will first discuss the criteria she employs and then the role of examples in her work.

Identifying Criteria

In terms of criteria that guide her work, Student Q invokes a criterion of **Audience** in considering what to include in her definition. After identifying two possible ways of defining threshold, she decides that both are important to include because "the person reading it would be like 'what about the other one?'" if only one was included.

Although she does not cite an explicit criterion to guide her, I will explore a possible criterion Student Q might be using in her move to add an additional phrase to her definition to make a distinction between two ideas. She notes that *threshold* could stand for "some point that is built up to" and *breaking point* could be the word signifying "just a barrier." Why does she feel the need to use a different word for each of these ideas and separate them out?

It isn't clear whether, if pressed to explain why she wanted to disambiguate these two ideas, she could articulate a clear criterion guiding her work. Perhaps she has some kind of criterion that the definition should be **Conceptually Bounded**, i.e. that one should separate ideas and use more specific naming if ideas seem different enough. The data doesn't allow us to say. Many people would likely have an intuition that definitions aren't supposed to be too long and **Conceptually Bounded** might help in managing this, though it necessitates that the designer bring in other criteria to decide how general or specific the definition needs to be. One could imagine a design contexts, perhaps in physics research, where there might be a reason to include only the idea of a *breaking point* in the definition. In Student Q's case, she does not see any clear reason to include one idea over the other, noting that threshold is "...one of those things that not everyone is going to 100% agree, because everyone is going to have slightly different interpretations..." In connection to this, she expresses some intuition that certain examples can be difficult to categorize, citing tomatoes as an instance of confusion as to whether they should be considered a fruit or vegetable. The tomato debate is an interesting one as it highlights different definitional design constraints in different communities of practice. As the Oxford Dictionary Online explains, the confusion arises from "differences in usage between scientists and cooks" (Oxford Dictionaries, n.d.). They go onto note that:

> **Scientifically speaking**, a tomato is definitely a fruit. True fruits are developed from the ovary in the base of the flower, and contain the seeds of the plant...As far as cooking is concerned, some things which are strictly fruits, such as tomatoes or bean pods, may be called 'vegetables' because they are used in savoury rather than sweet cooking. The term 'vegetable' is more generally used of other edible parts of plants, such as cabbage leaves, celery stalks, and potato tubers, which are not strictly the fruit of the plant from which they come. Occasionally the term 'fruit' may be used to refer to a part of a plant which is not a fruit, but which is used in sweet cooking: rhubarb, for
example...So, the answer to the question is that a tomato is technically the fruit of the tomato plant, but it’s used as a vegetable in cooking.

Thus, Student Q notes a possible conceptual disambiguation she could make, separating threshold into threshold and breaking point, however, she keeps both ideas in the current form of her crafted definition. It appears that she cannot find an additional criterion that might suggest she should separate them out more explicitly. Indeed, at the very end of the workshops when she was asked whether anything else should be added to the definition, she wrote, “something on how to define a point where people can agree on the division between 2 categories, even if its an arbitrary division.”

The Interplay Between Examples and Definitional Words and Phrases

In this episode there is additional evidence of participants employing examples to clarify and check new words or phrases in the definition, something that was seen in Episodes One and Two. Specifically, Student Q tried out her new build up phrasing with an example of crossing through a doorway. This helped her to question the necessity of a build up: “If you’re talking about a threshold being a door, is there a build-up to the door?” The door example was one she had brainstormed early on so it likely that she considered it to be a core one. Another place where a new phrase is checked with an example is when Student Q wrote down the new idea of two ways to see threshold, “some point that is built up to, or just a barrier.” She immediately followed this with two examples: building up strain on a pencil until it breaks, and sneezing, which, from her perspective, lacked a build up. Specifically, she said, “Snapping a pencil has a build up of strain. A sneeze is just not not sneezing.” She appears to want to clarify her phrasing and attempts to ground it in representative examples. Snapping a pencil was one of the initial examples introduced by the interviewer at the beginning of the activity, so she may have relied on it as a core example. Since she wrote down the phrase and examples without narrating what exactly she was doing, I do not have enough evidence to say how and why she brings in the pencil example specifically.

Conclusion

In this episode we made a strong case for the criterion Consider Its Audience and a weaker case for the criterion Conceptually Bounded, though Student Q did demonstrate some intuition around this criterion in her move to put different words to different ideas in the definition. We saw how the consideration of an example, sneezing, in which the participant did not have a strong experience of a build up, eventually resulted in definitional refinement. We also saw the critical role of examples in clarifying and challenging definitional phrasing.

---

8 Is there really a build up to the door? I would argue that, if one considers a parameter involved in the threshold, perhaps some distance coordinate, there will always be some minute change of that parameter leading up to crossing the door threshold. In that sense, there could be something characterized as a build up. However, given that the purpose of the design task is to help younger people identify threshold in the world, I believe it to be quite reasonable to sort threshold examples into ones with an experiential feeling of build up and ones without, at least in some initial description.
Conclusion

In this chapter I presented three episodes of definitional refinement, identified and named criteria, noted practices associated with meeting criteria, and noted example use. Now, I will return to my framework to highlight its use in helping understand episodes of definitional refinement and discuss in more detail the set of criteria that I identified.

Framework

My framework for definitional refinement included examples, the crafted definition, DUK, and criteria. These elements interacted in different ways during various moments and phases of definitional refinement as shown in Figure 12, which is reproduced, below:

![Figure 12. Final Framework for Definitional Refinement. D1 represents an initial definition and D2 represents a refined one. Framework elements such as examples, DUK, the crafted definition, and criteria, are part of each of the blue and green boxes are sometimes hidden within the term “practices.”](image)

For each episode, I identified D1, the initial definition and D2, the final definition. For the first two group work episodes, there was a clearly delineated pre-phase and post-phase and I broke up the transcript to show these two phases. In the last episode, based on a one-on-one interview, the interviewer-participant interaction somewhat blended these two phases which tend to be distinct when a group is working on its own. Finally, I showed how a misalignment between two elements was always at the beginning of an episode of definitional refinement and made it possible. Each episode demonstrated a different possible misalignment shown in the above figure: Example/Definition, Criterion/Definition, and Example/DUK.
Summary of Criteria And Discussion

In my analysis I focused on naming and identifying criteria. To summarize the analysis of the criteria, I had a set of criteria that were identified through students’ explicit phrasing and arguments and/or history of work. I considered this strong evidence for the existence of students employing these criteria. I also identified criteria through students’ actions. In this case, I did not have enough evidence to say whether students had some intuition for the criterion or, if pressed, could explicitly argue for that criterion. I considered this weak evidence for the existence of students employing these criteria. I summarize all of the criteria that I identified in this chapter, below. Recall that criteria were given a name that would complete the sentence “The definition should...”

Strong Evidence Criteria

**Encompass Prototypical/Core Examples:** The definition should always cover examples deemed (implicitly or explicitly) to be most core or prototypical. Core-checking is the related practice of checking any changes to the definition against one or more core examples.

**Cover All Known Examples:** the definition should cover all known examples of the phenomenon being defined.

**Use The Simplest Words Possible:** the definition should use simple and straightforward words whenever possible.

**Cover the Conceptual Core:** the definition should include all of the important features or aspects that are core to the phenomenon.

**Meet The Design Purpose:** the definition should meet the purpose of the design activity, in this case, helping people identify threshold out in the world.

**Sufficiently Precise As To Not Be Mistaken For Another Word:** the definition should not be easily mistaken for another word. For example, consider the following definition for *bird*: “something that flies.” This definition could be mistaken for the word *airplane*. Definitions need enough precision to disambiguate the idea being defined from other near-by ideas.

**Consider Its Audience:** The definition should consider its audience. This might involve clarifying things that the audience might wonder about or have confusion with. This might also involve the use of words that the audience will likely understand.

Weak Evidence Criteria

**Disregard Outlier Cases:** if one example has a difference that is true of only a small number of the total number of accepted examples, then the definition does not need to change.

**Conceptually Bounded:** the definition should not include too many ideas. One should separate ideas out and use more specific naming if they seem different enough. This criterion will often need additional criteria to help in deciding how to characterize “different enough,” such the Meet The Design Purpose criterion.

Criteria Discussion

Now that I have identified a number of criteria, I will examine some notable similarities and differences across them.
One difference between criteria was that some became the overall goal of the refinement phase (e.g. **Sufficiently Precise As To Not Be Mistaken For Another Word** for Episode Two) with other criteria helping the group navigate making changes to the definition toward the top-level goal (e.g. **Consider Its Audience**). Theoretically, any criterion could be in misalignment with the definition and become the overall goal of the refinement phase if a participant notices this.

Participants had to navigate issues of generality, specificity, conciseness, and precision, which often pulled them in different directions. One can see this in the types of criteria they employed. First, there were two criteria that suggested the bare minimum of what the definition should include: **Cover the Conceptual Core** and **Encompass Prototypical/Core Examples**. Recall that the Conceptual Core criterion arose when one participant was attempting to get rid of “change” in the group’s definition during attempt at using the **Simplest Words Possible**. His group mate responded defensively that “change” was core: “that’s the whole threshold thing.” Thus, attempts at meeting any other criteria were subject to these baseline ones. Conciseness might be helped by the criteria of **Conceptually Bounded** and **Disregard Outlier Cases**. Additionally, in the introduction, I noted a research study where undergraduates employed the criterion of “Economy” which involved a student arguing to remove a phrase from the crafted definition because it could be derived from other information found there. This would also be a criterion that would support a definition in being concise. On the other hand, the criterion that definitions must be **Sufficiently Precise As To Not Be Mistaken For Another Word**, might require adding more words or phrases to the definition to support precision.

**Practices Associated With Examples**

In this chapter I noted many places where students used examples for various goals. An interesting new practice that I identified was “core-checking,” which involved trying out new words or phrases with a small number of the most core or prototypical examples. This was especially helpful when phrases were new or a big change had just been made to the definition. Core-checking is complementary to the use of boundary cases. Core-checking helps participants get some grounding in the basics of what they are defining before they subject their crafted definitions to boundary examples which might call aspects of it into question.
Chapter 4 – A Cross-Case Analysis of How Participants Make Sense of New candidate examples

Introduction

In the last chapter I built a framework for definitional refinement that involved criteria, definitionally unarticulated knowledge (DUK), examples, and the current state of the crafted definition. I focused on identifying criteria and naming some of the practices involved in definitional refinement. In this chapter I take on a different focus: how participants make sense of new candidate examples and the extent to which they use their crafted definition in doing so. These new candidate examples are either brainstormed by the students themselves as part of working on their design task for the book or they are presented by the facilitator/interviewer.

To shed light on this process, I follow two one-on-one participants and one group as they make sense of the same example in three different ways: accepting it, rejecting it, or deciding to name a new subcategory. I will then make an additional comparison of the process of naming a new subcategory by bringing in data from a different participant who also named a new subcategory. These two different cross-case comparisons will comprise the two sections of this chapter (see Figure 13, below).

![Chapter 4 Cross-Case Comparison](image)

*Figure 13. The cross-case comparisons made in the two sections of this chapter.*

Each section will have its own particular set of research questions. In section one, the bulk of my analysis will focus on the following questions: “To what extent do participants use their definition versus DUK in making sense of the new candidate example?” and “What are the practices associated with examples as part...
of their sense-making work?” I will also comment on any other of my framework elements that appear in the data and note any differences between groups that may have impacted participants toward different results. Section two will focus on the question, “What criteria are involved in the process of participants’ deciding that naming a new subcategory is warranted?” and “How and why do participants brainstorm additional examples to put into their newly articulated subcategory?”

By the end of this chapter we will be able to reach four conclusions:

1. Participants used their crafted definition or Similar DUK in making sense of these particular new candidate examples and did not bring in any Implied or Outside DUK.
2. By paying attention to the difference between DUK and crafted definition use, some heuristics emerge to support definitional refinement.
3. Participants demonstrated many competences with respect to examining, brainstorming, and judging examples.
4. Subtle differences in framework elements can lead to quite different results when students make sense of a new candidate example.

Research Questions

Recall from my introduction that this chapter will focus on the following questions: “To what extent do participants use their definition versus DUK in making sense of the new candidate example?” and “What are the practices associated with the use of examples as part of their sense-making work?” I will also comment on any other framework elements that appear in the data and note any differences between groups that may have impacted participants toward different results.

Methodological Details about DUK

One of the major analysis strands in this chapter involves the question of how much participants use their crafted definition versus DUK in making sense of a new candidate example. Speaking to this question requires precision in identifying these two constructs. I explain my classification methodology, below.

Identifying Use of Crafted Definition in Transcript Data

A word or phrase in the transcript is categorized as using the crafted definition if it uses the exact word or phrase from the current form of participants’ crafted definition. Minor syntactic variations are discounted. For example, change of verb tense will be ignored. For instance, if participants’ definition involved the word “crossing” and they use the word “crossed,” this would be counted. Similarly, if an adjective in the definition had an “-ly” construction such as “quickly” and a participant then used it without its –ly ending such as “quick,” (or vice versa) this would still be categorized under use of the crafted definition.

Identifying Types of Definitionally Unarticulated Knowledge (DUK)

I draw a hard methodological line between the use of words actually present in the definition and any other words participants use in interpreting new candidate examples. I label the latter DUK and will argue that differences in words that can appear subtle are worth naming and paying attention to. DUK will lie along a
spectrum: some DUK will closely resemble words in the definition; other DUK will lie completely outside of it. If we imagine a bulls-eye with the crafted definition in the center, the first layer outside of this would be synonym, super-, or sub-ordinate words (“Similar” DUK). Another layer out would be ideas and their associated words that are implied by the words present in the definition (“Implied” DUK). The outer-most layer would include words and phrases neither implied nor specified by the crafted definition (“Outside” DUK).

In my analysis of Outside DUK I will also look back at the history of participants’ work to if it is completely new with respect the workshop. If Outside DUK did come up before I will attempt to make claims about why it was not included into the crafted definition.

Implied DUK

Implied DUK is associated with general criteria that can be difficult to balance in crafting definitions: include as much useful information as possible and be transparent, but make the definition as compact and economical as possible. Identifying DUK as “implied” is somewhat tricky methodologically since it begets a further question: who is judging whether it is implied or not? The researcher? The participants themselves? The purpose of the present analysis is to understand how my participants are using the crafted definition toward classifying examples. This suggests that it matters whether the participants themselves would judge that something is implied by the definition. If they are aware of certain DUK being implied, then the definition is doing its job: helping them bring certain ideas to bear in classifying new candidate examples.

Another important question in evaluating DUK is whether the researcher can suggest any possible places for further definitional refinement. While it is important to understand whether participants would judge something to be implied by the definition, the researcher can also examine how central Implied DUK is to participants’ classification process. For instance, consider a situation where Implied DUK is central to how participants classify and is also used across many new
candidate examples in classification. This raises the question of whether it might be useful for participants to have a careful conversation about whether definitional refinement is warranted to include something so central. If participants will be using the definition in the future, refining it could guard against participants forgetting about important implications. If other people will be using the definition, it could also guard against other users not understanding implied inferences. It is also possible that an implication could be fairly obvious to most participants and not necessitate definitional refinement.

**Similar DUK**

I use the term “Similar DUK” to refer to the use of synonyms or super- or subordinate words. While Similar DUK use by participants can appear subtle and inconsequential, it can sometimes give us insight into important tacit influences on participants’ categorization process. Similar DUK analysis can also show places to work on balancing specificity and generality.

A hypothetical case can support this point. Let us examine two verbs that could be used interchangeably but convey slightly different meanings. Let’s imagine a simple hypothetical crafted definition: “When an agent causes a patient to Z.” Then, let’s imagine a participant classifying a new candidate example using this definition: “This new candidate example counts. The agent forced the patient to Z.” The participant seems to be using the verb “to force” interchangeably with the verb “to cause.” However, these verbs convey a different relationship between the agent and patient with “to force” implying that the agent caused the patient “to Z” with some resistance from the patient that the agent overcomes.9 “To cause” might imply to some that the patient experiences no such resistance. Thus, one question for participants involved in crafting this hypothetical definition might be: Do you want your definition to encompass all the possible verbs under the umbrella of “to cause” or do you mean a more specific subset of instances with a particular power dynamic between the patient and agent? These can be incredibly important details to determine for some contexts.

I use the broader phrase “Similar DUK” and not “Synonym DUK” as I don’t want to limit this construct to words that a thesaurus specifies only. For instance, if one consider what a standard thesaurus (*Merriam-Webster Online Dictionary*, n.d.) lists as the synonyms for “cause,” they list:

*Beget, breed, bring, bring about, bring on, catalyze, cause, create, do, draw on, effectuate, engender, generate, induce, invoke, make, occasion, produce, prompt, result (in), spawn, translate (into), work, yield*

There are words one could judge as quite similar to “causing” that we do not see here. For instance, if you “push” something over a boundary, one could judge

9 Talmy (Talmy, 1998) goes into much more detail on differences in causation-related language. He sets about to take “‘causing’ into finer primitives and set it naturally within a framework that also includes ‘letting,’ ‘hindering,’ ‘helping,’ and still further notions not normally considered in the same context.”
this as “causing” something to go over a boundary. The words are related, but not specified in a thesaurus. This particular case suggests that words may be related through being super-ordinate or sub-ordinate to one another: pushing is one type of causing. Thus, Similar DUK includes words specified by a thesaurus as well as super- or sub-ordinate words.

Outside DUK

It is possible for participants to bring up knowledge entirely outside of the definition that is neither implied nor specified by it. Analysis of Outside DUK can suggest important places for possible definitional refinement. Sometimes Outside DUK can be completely new to the workshop. If it is not new and was considered at some earlier point during the history of the workshop, examining this history can help the researcher make claims about why it was not included in the definition.

Section One: Cross-Case Comparison Of Stoplight Example

Focal Example: The Stoplight

In this section, I will show three different results of participants’ consideration of a new candidate example: the stoplight. The stoplight example is the following: imagine a group of cars sitting at an intersection. When the light turns from red to green, the cars go from being stopped to moving. A student in Group A originally brought up this example while her group was brainstorming examples to include in their book. Her group had an involved discussion about it that resulted in their naming a new subcategory.10 I was curious whether my one-on-one participants would name a similar subcategory or not so I asked them about this example during their interviews.

Presentation of the Data and Analysis

For each of the three stoplight example consideration episodes, data and analysis will be organized according to the same format. I will first provide a narrative of the data and then the transcript. I will follow this with the analysis in three parts: first, a comparison of definition versus DUK use, second, an examination of how participants engage in examining examples, and third, a comment on other framework elements seen in the data. I will compare across groups when appropriate. I might also bring in excerpts of data from the history of participants’ work if it helps in making sense of participants’ actions.

I will first examine Group A, the group whose work results in naming a new subcategory. As we continue on to the participants who accept and reject the example, we will then compare to Group A to see if we can say anything about what elemental differences might have led to the overall difference of rejecting or accepting the example instead of naming a new subcategory.

---

10 Some readers might wonder about the fit of the stoplight example as within the category of threshold. In terms of threshold-like phenomena, it is not the most prototypical. Indeed, we will soon see another participant reject this example. However, each group is working on categorizing according to their own definition and purpose so I will not pass judgment on the fit of examples out of that context.
**Group A Decides To Name A New Subcategory: Data and Analysis**

Recall that Group A consists of students M, L, and N, as shown in Figure 2, reproduced, below.

![Figure 2. Group A included Students M, L, and N, clockwise from left](image)

**Episode Summary and Transcript**

This episode is roughly five minutes in length and occurs forty minutes into the workshop. Leading up to this episode, the group was engaged in brainstorming new candidate examples that kids might be particularly interested in and were talking about games. The game “red light, green light” was brought up. One participant, Student L, connected the game to the real experience of cars at stoplights and then it became a contentious example. Student N saw this example as different and articulated this difference with the help of the definition. Since the group’s definition involved the idea of boundaries, Student N argued that there seemed to be two boundaries, something that they hadn’t seen in other examples before.

At this point in time, the group’s note-taker had written down the following definition with a small grammar issue: “A threshold is a boundary. Forces help you cross that boundary. When you cross that boundary, different things start to happening than what was previously happening. (Where things change.)” The episode’s transcript follows:

Student L: *That’s a really simple one too, with the red light vs. the green light*
Student N: *Oh! Red light, green light*
Student L: *I think that should be included*
Student N: *Oh my gosh! Kid’s games! How fast do you run in red rover to actually get to the other side? How hard do you have to hit them?*
Student L: *[talking aloud what she's writing] Red Light, Green Light. I think that's like a pretty good like*
Student N: *And we could be like*
Student L: *Yeah, like, when it changes, then the cars move*
Student N: *Oh but that’s bringing in like, the threshold, well, what’s the threshold there? And what’s the force? And what’s changing? ’Cuz it's like, the light is causing the*
cars to move so is it the cars crossing the stopping/go ing boundary or is it the light crossing the green/red boundary?
Student L: There's the light changing, yeah
Student N: Oooh. Uh oh. Some things have to change to cause other change.
Student L: That's another
Student N: That's another thing entirely. Should we bring that in? 'Cause that's another thing that they'll [the kids, the audience of the book, will] be like 'well, is this a threshold?' Okay, so
Student L: Yeah we should figure that out
Student M: There's multiple thresholds involved
Student N: Mm, so like, systems of thresholds also exist
[The group goes on to articulate this as a new subcategory and brainstorms other examples that would also fit within it. This process will be explored in the next section]

In this episode, Student N does most of the work in interpreting the new candidate example. She uses aspects of her definition as well as DUK.

Analysis

Definition Use
First, let us identify definitional phrasing that is used in making sense of the new candidate example. The current version of the definition that the group had written down is, “A threshold is a boundary. Forces help you cross that boundary. When you cross that boundary, different things start to happening than what was previously happening. (Where things change.)” Student N does all of the talking in making sense of this example and if we examine the transcript, she references five words that are part of the definition: threshold, force, changing, crossing, and boundary. These terms are bolded in her transcript, below:

Student N: Oh but that’s bringing in like, the threshold, well, what’s the threshold there? And what’s the force? And what’s changing? ’Cuz it’s, like, the light is causing the cars to move so is it the cars crossing the stopping/go ing boundary or is it the light crossing the green/red boundary?

Student N: Oooh. Uh oh. Some things have to change to cause other change.

Student N uses almost all of her definitional phrasing in parsing through the new candidate example and it helps her to name that the example involves two different boundaries and two different changes. She also specifically names the “different things [that] start happening” and “what was previously happening” on either side of the boundary: “stopping/go ing” and “green/red.”

Similar DUK Use
Next, let’s examine her transcript for DUK. We see one word that is not part of the definition: causing. This term is bolded in her two transcript excerpts, below:
Student N: Oh but that’s bringing in like, the threshold, well, what’s the threshold there? And what’s the force? And what’s changing? ’Cuz it’s like, the light is causing the cars to move so is it the cars crossing the stopping/going boundary or is it the light crossing the green/red boundary?

Student N: Oooh. Uh oh. Some things have to change to cause other change.

The use of the word causing is Similar DUK. The participants’ crafted definition states that, “Forces help you cross that boundary.” Causing and helping are related verbs, but helping implies a support: either removing a blockage to an agent achieving a result, or helping an agent to achieve that result more quickly. Consider a hypothetical sentence. If I “helped Student Z pass the exam through tutoring,” I might share in some of the “cause” of Student Z passing the exam, but Student Z also played a (perhaps greater) role. The only cases in which “to cause” might be interchangeable with “to help” would be ones in which there were equal forces that cause a threshold event to occur. For instance, consider two people applying equal forces in pushing a block over the edge of the table. Each of the individual people both “helped” and “caused” the block to fall. Very few (if any) of the examples Group A had considered had multiple equal forces, however.

Perhaps not surprisingly, then, Student N breaks out of this verbiage when parsing the new stoplight example. Neither Student N nor her group make note of this difference between the crafted definition verb use and her verb use. I did not see any evidence of participants paying attention to definitional differences in verbs or adjectives in my entire data set. It might be the case that my participants are unaware of this particular heuristic. It might also be the case that my participants judged this heuristic unnecessary according to the level of precision of the design goal or had other tasks to work on that were deemed more important. There is not enough evidence to say.

Analysis: Other Elements that Contributed to Group A’s Process

Example Consideration Earlier in Workshop

An additional factor that likely played into Student N’s articulation of a new subcategory was that the group had discussed a very similar example of a block starting to move during an episode of definitional refinement about twenty minutes prior to this episode. If Student N had not considered the cars crossing the stopping/going boundary as a threshold example, there would not have been a “system of thresholds” to even consider. In the following excerpt, we see Student N deciding that a block going from “moving to not moving” involves a threshold. This piece of transcript is part of an episode of definitional refinement explored in Chapter 3 where the group is trying to disambiguate forces from thresholds. One group member asks, “what is a threshold?” and this excerpt is part of the dialogue that ensues:

Student N: [a threshold is] when it transforms into something else, I guess, but that’s still, like, sort of, like, if you push something, it transforms from stationary block to a
moving block, but it's not a threshold until it goes over a cliff, or it, like, you push it too fast and it melts
Student M: It's also, like, a threshold, the fact that it goes from like, not moving to moving, is threshold
Student N: Oh that's true, that's true
Student L: Ohhhh

In this excerpt we see that the group had considered what type of motion counted as threshold and specifically decided that going from “moving to not moving” involved a threshold.

Criteria
Lastly, we also see a criterion at work in the decision to continue thinking about this example. In this case, the criterion is not about what “a definition should be,” but instead, what sub-categories should be noted in the book. Student N asks, “Should we bring that in?” and makes an argument invoking the criterion that considers their Audience, highlighting that the children reading the book will likely wonder about this example: ‘Cause that’s another thing that they’ll be like 'well, is this a threshold?’” Student L agrees with this argument: “Yeah we should figure that out.” They go on to articulate and populate the category of “systems of thresholds” which we will explore in detail in the next section.

Group A Analysis Summary
We have seen how the definition, DUK, and one criterion played a role in the participants making sense of a new candidate example. Although Student N used phrases in the definition in her work, she also used a verb different from the one articulated in the definition. A criterion was then invoked in the group’s decision to more formally name the subcategory.

Student Q Accepts A Version Of The Stoplight Example: Data And Analysis

Figure 5. Student Q, on the right; Interviewer on the left

Episode Summary and Transcript
We now move on to Student Q. Recall that Student Q is the student in Figure 5, reproduced above. We will see that at first she interprets the stoplight example in light of some of her real life experience of driving. Once the interviewer puts a further constraint on the context of the example, that it occurs in a small town where people follow traffic laws, Student Q immediately agrees that the example
fits. She makes a few different types of arguments in doing so, connecting the stoplight example to another example she had talked through in-depth previously in the interview: crossing through a doorway. At the end of the transcript provided, Student Q goes on to place the example within a sub-definition she had previously defined. We do not provide this part of the transcript as it would require going into detail around Student Q's previous sub-definitional creation and it is not necessary for the analysis points we are making at this time. We begin the transcript at the moment where the interviewer introduces the new stoplight example:

Interviewer: *Like, a traffic signal, the light going from red to green, and the cars going from stopped to go*

Student Q: *I don’t really see that one, I think it’s 'cuz mostly, people are kind of crazy drivers and red doesn’t always mean stop, red sometimes means gun it*

Interviewer: *If we imagine some small town*

Student Q: *Where people actually follow traffic laws?*

Interviewer: *Where people actually follow traffic laws, yeah. Would that count?*

Student Q: *Yeah, yeah, yeah 'cuz there’s this point where you’re told you can go and there’s this point where you’re told you have to stop. This is assuming there would be no yellow lights 'cuz that never made sense*

Interviewer: *Yeah, maybe it’s a flashing*

Student Q: *It’s like the sign warning you about the door -- it’s about to turn red so you know, in preschool they tell you to slow down and when your mom’s teaching you how to drive she just tells you to gun it, make the light*

Interviewer: *Just do it anyway, okay, but in the scenario that the light turns from red to green and the people go from stopping to then they go, then that’s a pretty good example?*

Student Q: *Yeah*

[Student Q then makes a move to place the example within a subcategory of threshold that she previously named.]

Analysis

Definition Use

Student Q uses some phrases found in her definition in making sense of the new candidate definition. I will bold words in the definition that she uses. Student Q's crafted definition was the following:

“A finite **point where** something turns into something else, or the **point** at which something happens when you act on an object. **Two ways of seeing threshold: some point that is built up to, or just a barrier...**”

Student Q makes the following statement in agreeing strongly (“yeah, yeah, yeah”) that the small town stoplight example counts: “’cuz there’s this **point where you’re told you can go and there’s this point where you’re told you have to stop.”*

Indeed, her explanation does appear closely related to her definition: both reference a particular point and she also specifies the something ("you’re told you can go") as well as the something else ("you’re told you have to stop").
There is some evidence that Student Q might also be paying attention to the phrasing of “finite” in her definition. She articulates one qualifying assumption in accepting the small town stoplight example: “This is assuming there would be no yellow lights ’cuz that never made sense.” One possible interpretation of this statement is that she considered the adjective “finite” in her definition and this led her to notice that yellow lights affect whether the point between being told you can go or stop is instantaneous. Another possible interpretation would be that her personal experience of her mother telling her to “gun it” is underlying her statement about yellow lights. Indeed, her phrasing “that never made sense” suggests a personal history around yellow lights.

I will bring in two short excerpts from Student Q’s interview history to give additional evidence for the instantaneous interpretation of Student Q’s yellow light qualifying assumption. Both excerpts occur earlier on in the workshop as the interviewer is asking Student Q about various new candidate examples. In the first excerpt we will see Student Q use almost the exact same phrasing in interpreting a new candidate example. In the second excerpt she will reject an additional new candidate example that does not happen instantaneously but instead is more of a “gradient.”

One of the first new candidate examples the interviewer presented to Student Q after she crafted her definition was about milk going bad in the fridge. Student Q immediately accepted the example with the following phrasing: “Yeah! ’cuz I mean, There’s a point where it’s not bad and there’s a point where it is.” This is the exact same phrasing she used in interpreting the stoplight. In this case she also specifies the something (“where it’s not bad”) as well as the something else (“it is [bad]”).

While I might challenge Student Q on her interpretation that there is an instantaneous point between milk being good or bad, she does not notice any problem with the fit at this time. Her interpretation of the milk example contrasts with her interpretation of another new candidate example, ‘getting old.’ She ultimately rejects the example of ‘getting old’ because it does not happen instantaneously: “it’s more of a gradient than, like, a point.” She also makes a comparison to another accepted example in backing up this statement. Her transcript is below:

Student Q: There’s no concrete place where it happens I feel. It’s, like, a process, so like, when you boil water, there is a temperature at which it becomes vapor. There’s no second, your 43rd second of your blankety-blank year is not the point at which you get old
Interviewer: Right so, it's a little hard to say the point. So, in your opinion, should it--
Student Q: It feels like it's too, what's the word, like it's more of a gradient than, like, a point.

In Student Q’s discussion of ‘getting old,’ she articulates that there is no exact point where a change happens, that it is more of a gradient. This is further evidence supporting an interpretation that Student Q made a qualifying assumption about yellow lights due to her noticing that it might affect the “finite” nature of the point of change. The evidence is not entirely conclusive; it is still possible that she was
uncomfortable with yellow lights due to her personal history. The evidence does suggest, at least, that Student Q sometimes considers how spread out or instantaneous a change might be in classifying examples.

**DUK Use**

Student Q does not use any DUK in her making sense of the new candidate example.

**Example Use: Comparison To An Example From Prior Work on Refining The Definition**

Student Q also compares the yellow light of a stoplight to an example she had previously discussed, a doorway. This is not immediately obvious in the transcript since she does not explicitly call out the example by name, but instead uses the phrase, “It's like the sign warning you about the door.”

This is a reference to an example she had discussed during a prior episode of definitional refinement: ‘crossing through a doorway.’ The doorway example played a big part (see Chapter 3 for more detail) in her process of questioning whether the idea of build-up was necessary to her definition. She ultimately decided that it was not necessary and added a clarifying phrase to her definition, that a threshold can be, “some point that is built up to, or just a barrier.” During this prior episode she had joked that a door does not have a sign warning you that you’re about to approach: “But at the same time, if you’re talking about a threshold being a door, is there a build-up to the door? I guess you’re walking toward the door, there’s not signs, like, 'there's a door here, get prepared!' -- you're about to have a door! So I dunno”

In the present episode, Student Q connects to the yellow light to the door warning sign: “It’s like the sign warning you about the door -- it's about to turn red so you know, in preschool they tell you to slow down...” It is difficult to conclusively argue why she connects to the stoplight example. She had decided that examples could have a build up or not and still fit under her definition, so she does not appear to be concerned about build up or lack thereof. One interpretation is that she simply notices a similarity between a sign warning you about a doorway and a yellow light warning you about an upcoming stop. She might also be making an argument against yellow lights. If doors do not typically have warning signs, perhaps stoplights do not need yellow lights either: “when your mom's teaching you how to drive she just tells you to gun it, make the light...” Thus, Student Q's connection to the doorway example could be considered any of the following: an articulation of an interesting similarity she noticed across examples, an attempt at making an argument against yellow lights, or some combination of the two.

**Practices Associated With Examples**

Student Q’s very first move in considering the stoplight example is to give a more particular version of it based on her personal experience. She acknowledges the intended purpose of stoplights, but suggests that this is not always followed: “I don't really see that one, I think it’s ’cuz mostly, people are kind of crazy drivers and red doesn’t always mean stop, red sometimes means gun it.” She appears to bring in a
criterion of **Typicality** when judging this version of the example. We see this in her use of the word “*mostly,*” that “*mostly, people are kind of crazy drivers.*” She also demonstrates an understanding that this example can be quite context-dependent when she compares pre-school to her mothers’ instructions: “*in preschool they tell you to slow down and when your mom’s teaching you how to drive she just tells you to gun it, make the light.*”

The facilitator then makes a move to have them consider a special version of the stoplight example, one in a small town where people follow traffic laws. Student Q readily accepts this example as counting as threshold, but then makes an additional clarifying move: “*This is assuming there would be no yellow lights ‘cuz that never made sense.*” This is a somewhat surprising shift for Student Q to consider acceptable. In the initial dialogue when the more general example was first introduced, Student Q made a move to categorize it according to what was most typical or what “*mostly*” occurred. One possible interpretation of her comfort with shifting the small town traffic light example away from what is most typical of stoplights – they do usually have a yellow light – is that she already considers the small town context to be a not-typical case. This could mean that further not-typical constraints on the example are deemed acceptable within an already not-typical context.

There is further evidence for this interpretation if we consider a move that Student Q performed with an entirely different example: sneezing. Student Q also articulates a more particular version of this example that is connected to her personal experience, but is stalled when she cannot judge its Typicality. More specifically, when the facilitator first asked her about the new candidate example of sneezing, she first proposes whether it might be similar to coughing but brings in some personal experience that problematizes this initial proposal: “*There’s a point at which you have to sneeze and let’s say it’s more like coughing. You can take a little bit of annoyance in your throat to the point where you’re, like, coughing up organs and stuff and there’s kind of like a trigger point, I dunno, my sneezes always catch me by surprise usually.*” We then see Student Q evaluate her personal experience with respect to typicality, deciding that she does not have enough information to make a typicality judgment: “*Sneezing seems to just come out of the blue, maybe my sneezes are weird, I don’t really know, I guess if I knew more about sneezing, what caused sneezing.*”

Student Q demonstrates a practice of specifying possible particular versions of a more generally presented new candidate example and evaluating them according to a criterion of **Typicality.**

**Student Q Analysis Summary**

In this analysis we saw how Student Q used her definition and no DUK in making sense of the new candidate example. She also demonstrated competence in considering more specific versions of the general example and judging them according to a criterion of Typicality. We also saw some that Student Q considers whether a transition point is instantaneous or more of a gradient in judging the fit of new candidate examples.
Student G Rejects The Stoplight Example: Data and Analysis

Figure 4. Student G, on the right; Interviewer on the left

Episode Summary and Transcript

In this episode, occurring about twenty minutes into the interview, we examine how Student G rejects the stoplight example. Recall that Student G is the student in Figure 4, reproduced above. Leading up to this episode, Student G had first crafted her definition: “There is a continuous action on an object, a limit is reached where the action has reached a maximum and causes a sudden reaction. This limit would be the threshold.” Next, the interviewer and Student G had a conversation about four new candidate examples that the interviewer introduced. In the process of discussing these other examples, Student G named one new subcategory, “gradual thresholds.” She decided that not all threshold examples needed to have a “sudden reaction” as she had specified in her definition. In naming this new subcategory, she had placed two examples within it: ‘milk going sour in the fridge’ and ‘the destruction of trees affecting an ecosystem.’ She had not made any moves to change her definition to remove the word “sudden”, however, only naming the new subcategory. She will reference this subcategory in her interpretation of the stoplight.

The stoplight episode is roughly two minutes in length. When the interviewer brings up the stoplight example, Student G’s immediate reaction is that it is not a good example. She then uses aspects of her definition in an attempt to articulate what is wrong with it. She also checks the example with her subcategory of gradual thresholds, deciding that it does not fit there either. The transcript follows:

Interviewer: If cars are stopped at a traffic light, the light turns from red to green and the cars go, is that, does that count? 
Student G: I wouldn’t say so. There’s a reaction to the light turning green, that is, the cars go, but I don’t see there, again there’s, the lights are timed, I know that, or there’s a certain way lights work so that it’s kind of a structural system to it, it’s not like it waits there until there’s like 5 cars and then it turns green, it’s kind of an automated process, so there’s no “maximum reached” and the number of cars waiting at a red light can vary before it turns into a green light. So I don’t see the continuous action or that maximum being able to be defined, um, and it’s not a gradual process either as in the sour milk or the trees ’cuz I think that’s just something that is a little more, I wouldn’t say it’s a threshold
Interviewer: Yeah, so it's just kinda automated, so the light turns at a certain time and the people go when the light turns green
Student G: Yeah it's a reaction
Interviewer: So there's a reaction but it's not really the
Student G: the maximum or the limit that's reached
Interviewer: And there's not really a continuous build up in some way
Student G: mm hmm
Interviewer: Yeah yeah that makes sense
Student G: 'cuz other times, I'm trying to think if there would be a continuous build up somewhere but not really, I mean sometimes you go through green lights without having to wait for a red light, but there's not that step by step process.
Interviewer: mm hmm
Student G: Yeah I wouldn't include it in the book, either.

Analysis

Definition Use

First let's explore how much of her response draws from her crafted definition. Her definition that had been crafted twenty minutes prior to this interaction is provided, below. I will bold the words that she references from it:

“There is a continuous action on an object, a limit is reached where the action has reached a maximum and causes a sudden reaction. This limit would be the threshold.”

Her initial response includes many elements of the definition as we see bolded, below:

Student G: I wouldn't say so. There's a reaction to the light turning green, that is, the cars go, but I don't see there, again there's, the lights are timed, I know that, or there's a certain way lights work so that it's kind of a structural system to it, it's not like it waits there until there's like 5 cars and then it turns green, it's kind of an automated process, so there's no "maximum reached" and the number of cars waiting at a red light can vary before it turns into a green light. So I don't see the continuous action or that maximum being able to be defined...

Student G's initial reaction is that the example does not fit. In justifying this position she first acknowledges that the example does have one aspect of her definition: a reaction. However, Student G bring in some knowledge about lights knowledge around the ways that lights work. Because of the “automation” with this process and the fact that the number of cars waiting at a light can vary before it turns, she proposes that neither the continuous action nor the maximum can be defined. Student G goes on to clarify:

Student G: Yeah it's a reaction
Interviewer: So there's a reaction but it's not really the
Student G: the maximum or the limit that's reached
Student G interprets the stoplight example entirely within her definitional phrasing until the interviewer introduces some Similar DUK.

**DUK Use**

The only DUK that Student G uses would qualify as Similar DUK. The interviewer introduces the phrase “continuous build up” into the conversation instead of the more general definitional phrase “continuous action” that Student G had been using up to this point. Student G takes up this “build up” phrasing (“I’m trying to think if there would be a continuous build up”) and then introduces the further phrase “step by step process.” The Similar DUK is bolded in the transcript, below:

Interviewer: *And there’s not really a continuous build up in some way*
Student G: *mm hmm*
Interviewer: *Yeah yeah that makes sense*
Student G: *’cuz other times, I’m trying to think if there would be a continuous build up somewhere but not really, I mean sometimes you go through green lights without having to wait for a red light, but there’s not that step by step process.*

When Similar DUK is used, I ask the question: what is the relationship between the similar phrasing and the definitional phrasing? Is there anything we can learn about how the definition could be improved?

One main question is whether the definitional phrasing is more general or more specific than the synonym phrasing, which can suggest places to refine the definition. In this case, “build up” fits under the more general definitional phrase of “action.” At first blush, this seems reasonable, though Student G could consider if a more specific phrase might be better according to her purposes.

Comparing her definitional phrase “continuous action” to “step by step process” does suggest one possible clarification she could make according to an **Audience** criterion. Although it is likely Student G would consider these phrases interchangeable, when I hear the word “continuous,” one way that I interpret it is in opposition to “discrete.” And yet, the phrase “step by step” suggests a set of discrete actions. What does Student G mean by “continuous action”?

I would argue that Student G means something like “a repeated action performed continuously,” rather than something like “not discrete.” In a prior conversation during the workshop, she discussed how threshold examples often involve, “continuous action, such as, like, poking or stretching...like blowing or tipping or pushing. There’s this continuous thrust or action being put on an object...which is the same thing that happens when you lean on a chair. You keep leaning and keep leaning and you reach a certain point where all of a sudden it tipped on the floor” This suggests that her use of continuous as a modifier does not mean that the repeated action cannot be discrete, just that it should be performed continuously. This interpretation of “continuous” would not be in opposition to the phrase, “step by step,” however, this may not be immediately clear to her potential audience. This would be a place where Student G could put more clarification into her phrasing.
Analysis: Practices Associated With Examples

Like Student Q, Student G engages in a practice of further specifying particular versions of a more generally presented new candidate example. Her specification of the automation of stoplights plays a big role in her rejection of the example. She also attempts additional versions of the stoplight example to see if she can find a fit, but ultimately rejects them as well, according to different criteria.

In interpreting the stoplight example, Student G immediately articulates two more particular versions of it: timed stoplights versus a stoplight waiting until there are a set number of cars to change. Like Student Q, she makes a typicality judgment: she "knows" that lights work in the former rather than the latter way: "the lights are timed, I know that, or there's a certain way lights work so that it's kind of a structural system to it, it's not like it waits there until there's like 5 cars and then it turns green."

Student G makes one additional attempt to come up with another version of the example that might fit: "'cuz other times, I'm trying to think if there would be a continuous build up somewhere but not really, I mean sometimes you go through green lights without having to wait for a red light, but there's not that step by step process." In this case it is not typicality that leads her to reject this version of the stoplight example since she acknowledges that it is sometimes the case that people do not have to wait for a red light. Instead, she judges this particular version of the example as not meeting her Similar DUK of a “continuous build up” and “step by step process.”

This is not the only episode in the data where Student G engages in specifying different versions of the same example and then judging their fit. It is instructive to consider one additional episode to give evidence that Student G’s engagement in this practice is not a fluke and is instead part of how she generally approaches making sense of new candidate examples.

Directly before the stoplight episode that we have examined here, the interviewer asked Student G about a different example, ‘turning twenty-one,’ a reference to the legal drinking age in the United States. The interviewer then had a minute and a half long conversation with her that resulted in her determination to not include it in the book.

In her first consideration of the example she specifies it as a legal distinction: "I think it's just something that the government picked because in other countries the legal drinking age is sixteen." She does not explicitly use definitional phrasing or DUK in her judgment, but implies that this would obviously not count.

She then proposes a different version of the example, the potential for destructive behavior: "Although I would say that turning twenty-one here in the U.S. is conducive to a lot of destructive behavior, so I guess in that sense there could be a threshold, but I think I'm just applying stuff to make it work." She then moves to entirely reject ‘turning twenty-one’ as an example.

How might we interpret her phrase, “I'm just applying stuff to make it work”? There are three main interpretations that are difficult to argue between conclusively. There is some chance that this is another instance of the criterion of
Typicality. It could be the case that this is a reference to the variable drinking ages across countries such that destructive behavior might not be typical in all countries. Within the U.S., however, she appears to judge destructive behavior as typical. We see this in her description: turning twenty-one is not "sometimes conducive" but rather "is conducive," to destructive behavior. This casts some doubt on the typicality interpretation. Another interpretive possibility is that the phrase represents her considering a criterion of Definitional/DUK Fit. It may be that even though destructive behavior is judged typical, Student G judges that there is a lack of build up in the government's somewhat arbitrary decision. Yet, this interpretation can also be cast into doubt when Student G says, "there could be a threshold."

A third possibility is that even though this particular example might be typical and could exhibit a threshold, Student G might be judging it as an implication of 'turning twenty-one' rather than representative of 'turning twenty-one' itself. Imagine if Student G had played out the consequences of the example further such that the destructive behavior caused Z which caused N that then caused K. If K is judged to include a threshold, could it qualify under the umbrella of 'turning twenty-one' or might it be its own new candidate example? Student G's wording of "applying stuff to make it work" could be interpreted as "tacking on consequences." The evidence is not conclusive but it is suggestive that she might have some intuition here. I name this criterion **Fidelity To The Original Example:** when judging a more particular version of an example, is it still representative of the general phrasing of that example? I can see two ways that students might invoke this criterion. This first would be like Student G: judging that the threshold exists within a string of consequences and is no longer representative. Another way might be judging that a particular example was zoomed too far in or out of such that it is no longer representative. For instance, imagine a hypothetical particular version of the stoplight focused on the electronics of stoplights and a threshold focused within a component of a circuit. This might also fail the fidelity criterion.

Ultimately, Student G has shown competence in naming multiple particular versions of a general example phrasing and judging their fit. We have seen strong evidence for two criteria in her judging these particular examples: **Definition/DUK Fit** and **Typicality.** There is inconclusive though suggestive evidence for one further criterion: **Fidelity To The Original Example.** "Typicality" and "Fidelity To The Original Example" show how students understand that examples like "sneezing" might be instantiated in different ways. Thus, **Definition/DUK Fit** must be supplemented by additional ways of judging various instantiations.

**Analysis: Use of Examples Related To A Named Sub-Cateogry**

The last thing we see Student G attempt before rejecting the stoplight example is placing the new candidate example within a subcategory she had named. Prior to this episode, Student G had created a subcategory of "gradual thresholds" in which she had placed two examples: "milk going bad" and "trees" in the sense of, "this gradual process of how many trees you can cut down before it affects the ecosystem." Student G decides that the example does not fit within her sub-definition either: "it's not a gradual process either as in the sour milk or the trees...I wouldn't say it's a threshold." Her subcategory of gradual processes was originally created when she considered the sour milk example and had an initial reaction that she did not
think it counted as a threshold. She then re-considered the milk example in light of her definition that led to the creation of her “gradual” subcategory.

She appears to check the fit of the stoplight example with this subcategory as a type of last resort move with reasoning along the lines of, “here’s a set of examples I thought did not fit at first; I should see if the stoplight could be similar to these.” This makes some sense; when she named this subcategory she decided that part of her definition was unnecessary but had not modified it. This would mean that an example could not quite fit her definition but still fall under her subcategory. Ultimately, she decides that it is not a gradual threshold either, and then makes the move to reject the stoplight example as within the category of threshold.

**Student G Analysis Summary**

In analyzing Student G’s rejection of the stoplight example we saw how she almost entirely used her definitional phrasing in interpreting it. She also had many competences in naming different particular versions of the stoplight example and judging their fit. Finally, she engaged in a practice of checking the example with her named subcategory.

**Comparison Analysis Across Group A, Student Q, and Student G**

One of the hardest arguments to make in this work is why participants went in these different directions that resulted in either naming a new subcategory, accepting, or rejecting the stoplight example. There were many differences in framework elements and their associated practices that could be interplaying: from the form of the crafted definition to the use of DUK to the bringing in of previously discussed examples. Participants also named different particular versions of the general stoplight example, often drawing on their personal experience or knowledge about more specific aspects of it. They also brought in different criterion in judging the fit of these examples. I will first look at differences in the definition and then other framework elements to see what is possible to say about why the various participants judged the stoplight example so differently.

First, let us consider the three definitions side-by-side in Table 2, below. I keep the words they used in considering the stoplight example bolded:
Participant | Definition
---|---
Group A | “A threshold is a boundary. Forces help you cross that boundary. When you cross that boundary, different things start to happening than what was previously happening. (Where things change.”

Student Q | A finite point where something turns into something else, or the point at which something happens when you act on an object.

Student G | There is a continuous action on an object, a limit is reached where the action has reached a maximum and causes a sudden reaction. This limit would be the threshold.”

| Table 2. All of the participants’ definitions with words bolded that they used in interpreting the stoplight example. |

Student G gives the most attention to naming what occurs during the pre-phase involved in threshold events: she notes that some kind of prolonged action occurs which reaches a maximum. Student Q and Group A have definitions more focused on the “boundary” or “barrier” itself with less of a description of what happens leading up to that point. Student Q even goes so far in her definition as to note that a build up is not necessary.

For Student G, one thing that we can say is that her specific knowledge about the automation of city stoplights was a big factor in her ultimate rejection of the stoplight example. None of the other participants explicitly brought this example-related knowledge into their considerations. This knowledge, in concert with words “build up,” “limit,” and “maximum,” in her definition, led her to reject the stoplight example. She made some attempts to consider different versions of the example that might fit, however, she did not uncover anything about the example to lead to its acceptance.

Student Q’s initial rejection of the stoplight example appears very similar to Student G’s. Student Q brought in specific example-related knowledge of her personal experience of people’s tendency to run red lights. She deemed this typical in the same way that Student G takes automation to be typical. Thus, one major difference between Students Q and G arose in that the interviewer found a version of the stoplight example, stoplights in small towns, where Student Q’s initial concern was no longer a problem. Student G would presumably not have accepted the small
town version of the stoplight example since she would likely have still considered it to be automated.

When the interviewer brings in the ‘small town’ version of the stoplight, Student Q does accept this version of the example. Her acceptance of the small town stoplight raises the question: why did she not see two thresholds in the small town case like Student N (of Group A) did? I would have expected Student Q and Student N to interpret this example in similar ways. Both have more of a focus on state change than on build up in their definitions, with Student Q explicitly noting that a build up was not necessary in her definition and Student N using the word “change” predominately in her consideration of the stoplight. Why does Student Q not “see” two thresholds in the small town stoplight example that she accepts? Student N immediately sees the stoplight’s change from red to green as a state change whereas Student Q does not seem to notice this possibility.

The different outcomes between Student Q and Student N are most likely caused by differences in the more specific versions of the stoplight that they considered. Perhaps, like Student G, Student Q has some specific example-related knowledge of the automation of stoplights or has some other underlying example-related knowledge that blocks her from seeing the stoplight as experiencing state change. Perhaps the fact that Student Q paid more attention than Student N to the yellow light in the stoplight blocked her from seeing it as simple state change. In fact, Student N did not bring up the yellow light at all during her consideration of the stoplight. While it is difficult to say conclusively, it seems quite likely that differences in the particular version of the stoplight example that Student N and Student Q focused on led to the their different outcomes. Student N considered the stoplight turning from green to red (which does not include a yellow light); Student Q considered the stoplight turning from red to green (which does includes a yellow light). It is interesting to note that something as small as whether a participant focused on the green to red or red to green transition of a stoplight could have been the difference between simply accepting an example or naming a new subcategory.

I have shown how there are clear differences in framework elements in participants’ consideration of new candidate examples. I have highlighted some of these differences and suggested ways in which they may have influenced participants toward different results.

**Conclusion to Section One**

In this first section of the chapter we examined three different outcomes from participants’ consideration of one newly introduced example, the stoplight. I analyzed and named the practices associated with my framework elements in participants’ considerations. We saw participants using only their crafted definition and Similar DUK in classifying new candidate examples and no Outside or Implied DUK. I showed competences in participants naming multiple specific versions of an example and judging their fit. This analysis gives us a baseline for what is important to participants’ making sense of new candidate examples and also gave us some heuristics that could be useful in definitional refinement more generally. The analysis also suggested how important the interplay between various framework elements are in participants’ decisions to ultimately craft a sub-definition. In comparing Group A to Students G and Q, we saw how, even if there are
commonalities across many of the participants’ framework elements, even one small element difference can impact them toward quite different results.

In my analysis of definition versus DUK use, students stayed surprisingly close to their crafted definitions. In prior work on definitions (Nachlieli and Sfard, 2003; Tall and Vinner, 1981), math students frequently employ knowledge outside of the definition given to them by a teacher. Many of my participants stayed almost entirely within their crafted definitional language in classifying new candidate examples and, if they used words outside of it, these words were close-by Similar DUK and not Outside DUK. This shows that crafting a definition from scratch can aid students in making the definition a useful tool for themselves. One caveat I should add is that DUK use in and of itself is not necessary bad if students are open to incorporating it into their definition. In fact DUK use, as I showed in Chapter Three, has the potential to support definitional refinement. Any time that a new candidate example has some surprising characteristic, DUK has the potential to come out. Since the stoplight episodes occurred over thirty minutes after participants had crafted their definition and they had each considered a diverse range of examples in the meantime, the lack of Outside DUK use, in the end, is not terribly surprising.

My DUK methodological discussion and analysis suggested some heuristics one could use more generally for definitional refinement. If one wants to refine their definition, one could similarly keep track of (perhaps record) the arguments one makes around trying to categorize examples and make note of the words used in comparison to the crafted definition. Similar DUK identification can often suggest places where the definitional language could be made more general or more specific. Implied DUK identification can suggest places where one might consider being more explicit in the definition. Outside DUK identification can often suggest big ideas that might be important to bring into the definition. In examining Similar DUK, we saw that there was some knowledge “hanging out” on the periphery of the definition that could be better articulated and incorporated into it. There are many criteria to balance, however, in doing so.

My analysis of the ways that participants named and judged more specific versions of the general example phrasing showed participants’ competences. First, they understood that examples could be context-dependent and often came up with multiple versions of the same example. They also brought many criteria to bear in judging these versions. One major criterion we saw involved participants’ moves to categorize an example according to typicality and even recognizing that their own personal experiences could be idiosyncratic. For instance, Student Q recognized that her own experience of sneezing needed to be evaluated against others’ and she needed more information in determining what was typical. I provided evidence for use of the following criteria in judging the fit of examples:

- **Typicality** (strong evidence)
- **Fidelity To The Original Example** (inconclusive evidence)
- **Definitional/DUK Fit** (strong evidence)

The first two of these criteria would supplement Definitional/DUK fit when students noted that particular versions of an example might be quite different.
Ultimately, this analysis gave a more specific set of practices associated with my framework elements and showed their interaction, shed light on participant competences, suggested that once students spend time crafting their own definition they do not stray far from it in making sense of new candidate examples, and suggested some heuristics for possible definitional refinement.

**Section Two: How Do Participants Name and Populate A New Subcategory?**

In the previous section we examined how participants make sense of new candidate examples and compared three different trajectories: acceptance, rejection, and a decision to work on naming a new subcategory. In this chapter we return to Group A where we left off: on the cusp of naming a new subcategory for systems of thresholds. We will examine how the group goes about crafting a new sub-definition, identifying framework elements and the practices associated with them. For comparison, we will also examine the path of one one-on-one interview participant, Student G, in her naming and populating a new subcategory of “gradual” thresholds.

This analysis will highlight the role of criteria in how participants decide whether naming a new sub-definition is warranted. It will also show how participants are not satisfied with naming a new subcategory with only one member. We will see participants engaging in brainstorming to populate the subcategory and examine how they do this.

I will present each episode in three parts: a summary of the episode, the transcript itself, and then the analysis of framework elements. I will start with Group A and then move on to Student G. For Group A, it is possible to gain additional insight by examining their written notes. However, this will require some explanation of the sub-categories they have already named and populated. I will provide this explanation and examination of their written notes after the transcript.

**Episode One: Group A Names and Populates “Systems” of Threshold**

**Summary, Transcript, and Explanation of Previously Named Sub-Categories**

This episode continues from where the first half of this chapter left off. Group A considered a new candidate example, the stoplight example, and Student N did most of the work in naming how the stoplight example was different from other examples they had considered. She articulated that it had two changes that occurred with one change causing the subsequent one: the stoplight changing color and the cars going from stopped to moving. Now we begin where the group then discusses whether this should be considered another subcategory as they have named two distinct sub-categories already. Student N advises the group to brainstorm more examples to make a determination. Group members then try two distinct ways of brainstorming examples that fit this new category. Ultimately, they agree on the fit of only one additional example: population dynamics within an ecosystem, with one species dying off causing another species to die off. The transcript is below:

Student N: *Oooh. Uh oh. Some things have to change to cause other change.*
Student L: *That's another*
Student N: That’s another thing entirely. Should we bring that in? ’cause that’s another thing that they’ll be like ‘well, is this a threshold?’ Okay, so
Student L: Yeah we should figure that out
Student M: There’s multiple thresholds involved
Student N: mm, so like, Systems of thresholds also exist
Student M: Is that another category?
Student N: I dunno
Student L: We could do that at the end [of the book], maybe, at the end of a certain--
Student N: I don’t know if that’s another category. What other systems of threshold can we think of?
Student M: Like, Ecosystems? I don’t know if that would be a threshold, it would be like, all these different organisms, affecting each other, a population of one, yeah...
Student N: If the population of rabbits increases, then this population of wolves increases until, the population of wolves is too high and then yeah
Student L: And they eat all the rabbits
Student N: That’s the classic example
Student L: And the wolves die. Yeah, we could do that, I think that’s, yeah, I think that’s a pretty good
Student N: Ecosystems, yeah. Okay, what about, in a game, or in a household, if you’re having kind of a bad day, a lot of things can compound that before you call your day bad, I guess
Student L: Oh, yeah. Huh.
Student N: What other systems? I’m having a hard time thinking of...
Student L: yeah
Student N: Systems of thresholds
Student M: Another system would be, like, wall street, like all these companies affecting each other but that’s like... [whole group laughter]
Student N: Yeah but I don’t know if we can be, like, “On wall street”... [tone suggests joking]
Student L: I was thinking dominoes, but, I mean, it’s like, all the same thing [gestures with her hand a repeated motion imitating dominoes falling over] -- just a bunch of thresholds
Student M: Dominoes is good, it’s like, a physical short one. Maybe we’ll think about more systems as we, like, contemplate this a bit more

Written Notes and Previously Named Sub-Categories

Two of the three participants put the idea of “systems of thresholds” down into the personal written notes they were taking during this episode. The third participant did not put any notes down about “systems of thresholds.” By examining these notes we will see that the participants struggled with where to place this new subcategory with respect to the other sub-categories they had named already.

To orient you to these notes, the students were keeping track of sub-categories of threshold toward making them into book chapters. The group’s professed goal around these sub-categories was to organize the chapters of the book in order of what was easiest for their audience to understand to what might be

---

11 What student M means by a “physical short one” will be explained in the next paragraph.
hardest. The group had made a distinction between social and physical thresholds, the former roughly having to do with people and usually lacking an exact physical mechanism. One example they had placed in this category was ‘how loud you can talk in a library before the librarian tells you to be quiet.’ They also made a distinction between short and long thresholds. This distinction was originally brought up around the examples of ‘aging’ and ‘water freezing’: that it takes time to transition from one state to another. During the workshop, the group had an additional clarifying conversation about this distinction and re-named these sub-categories to be sudden instead of short and continuous instead of long.

The group appeared to struggle with where to place the idea of “systems of threshold.” We see in the participant work below (see red arrows). One participant, Student L, wrote down the idea of systems outside of their matrix of social/physical and sudden/continuous. The other participant, Student M, made an attempt to put the idea into the matrix. See Figures 15 and 16, below.

Figure 15. Student L’s personal written notes. No attempt was made to put “systems of thresholds” within the social/physical and sudden/continuous matrix. Note: the red underlining of the sub-categories and the red arrow pointing to “systems” was added by the researcher.
Some of this struggle may have been around whether “systems of threshold” should be considered a subcategory or a super-category. The participants do not seem to differentiate between these two possibilities and use the general word “category” in discussing any of the distinctions they had made. Ultimately, the participants decide to include only the social/physical and sudden/continuous distinctions into their book given their time constraints. Near the end of the workshop, Students N and L tell the facilitator about the four chapters and their ordering and we see how they have dropped the idea of “systems of thresholds”:

Student N: Yeah, so those are our categories. We had, like, sudden and continuous and then social and physical, so we had, like, four categories and we were gonna do sudden physical, then sudden social, then continuous physical, then continuous social
Facilitator: Oh, cool, that seems reasonable.
Student L: Yeah, the ones you would notice more, like, immediately to the ones that take a little more time to think about.
Student N: Yeah, canaries to ostriches as it were

The group does not further consider the “systems of thresholds” idea except in passing as a possible last chapter for their book.

Analysis

Criteria

Let’s first identify the criteria in this episode. In the previous section of this chapter we identified one Audience criterion where Student N advocated for bringing the systems of threshold idea into the book since the readers might wonder
about it. Specifically she said, “Should we bring that in? ’cause that’s another thing that they’ll be like ’well, is this a threshold?’”

We also see Student N bringing in an additional criterion with respect to whether a new subcategory is warranted: “I don’t know if that’s another category. What other systems of threshold can we think of?” Her challenge to her group suggests that she has a criterion around subcategory formation that naming a new subcategory requires more than one example. It is not clear whether she generally thinks sub-categories must have more than one member or whether having more than one member would mean that the subcategory is more important, warranting its inclusion as a category in the book. Her words could be interpreted either way: “I don’t know if that’s another category” could mean exactly what it says, or, it could mean, “I don’t know if that’s another category [that is important enough for the book.]” There is not enough data to say conclusively what her reasoning is. In the former case, this criterion could be named Sub-Categories Have More Than One Member; in the latter case it would be Important Sub-Categories Have More Than One Member.

Whatever criterion might be underlying Student N’s move to populate the subcategory, this was a move that all of the participants in my data made any time the possibility of naming a new subcategory arose. Other participants may not have been as explicit as Student N in articulating the need to find other category members, however, any time they were considering naming a new subcategory, they found at least one more example to populate it.

It is not clear how many examples Student N might require to officially name a new subcategory, but it appears to be more than two. The group can only brainstorm one more example, ecosystems, that they decide fits the subcategory. Student N moves to close their work but does not imply that it is done. She suggests that they should return to it at some later point. This means that we do not have enough data to say how many examples Student N would judge as enough.

Practices in Brainstorming More Examples

The participants demonstrate two different approaches to brainstorming examples which both include narrowing the types of possible examples to consider. First, Student M employs a strategy of trying to brainstorm new candidate examples of “things that affect each other” to then examine these examples for the possibility of multiple thresholds. “Things that affect each other” is similar to the way the group has been discussing the new subcategory with some additional aspects removed (e.g. double changes are no longer mentioned). Thus, Student M is not brainstorming randomly, but within a narrowed space that comes from choosing one of the characteristics of how “systems of thresholds” had been loosely defined. Student M comes up with the example of ecosystems first. It is not immediately accepted by her group since she has only named “all these different organisms, affecting each other” and has not named the two thresholds. Her group members help her find places within the example that exhibit two connected thresholds: rabbits all dying causing the wolves to die as well. This is productive for the group: they find an example that they all agree fits in the new subcategory. I label this kind of narrowing practice “Characteristic Oriented Brainstorming.”
Student M finds an additional candidate example of companies affecting one another: “Another system would be, like, wall street, like all these companies affecting each other.” Instead of aiding her in finding two thresholds in this example as they did in the ecosystem case, the group erupts in laughter. It is most likely that the group is considering their audience of younger people in not pursuing this example further. Evidence for this lies in Student N’s joking tone when she says, “Yeah but I don’t know if we can be, like, ‘On wall street...’” and the group moves on without further pursing the wall street example. Although the group does not further pursue the example in this case, it seems likely that they could have found one threshold causing another within this example if they had spent some time on it. This type of Characteristic Oriented brainstorming is quite productive for the group. Students N and L engage in a different kind of brainstorming that involves narrowing unrelated to characteristics. Instead, they narrow to particular smaller sets of things: games or things that happen in a household. Student N comes up with the new candidate example of many things compounding to give you a bad day. Student L comes up with the new candidate example of dominoes. Both students find new candidate examples in these smaller sets to consider, but ultimately decide that they do not quite fit. I label this kind of narrowing practice “Subset Brainstorming.”

Student L comes close to finding another example within the games context. However, she exhibits a close attention to detail when she considers but then rejects the example of dominoes: “I was thinking dominoes but, I mean, it’s like, all the same thing [she gestures a repeated motion of dominoes falling over with her hand that suggests that each threshold looks the same], just a bunch of thresholds.” Although she has found an example of one threshold event causing the next threshold event, she makes a distinction that “systems of thresholds” should involve two different kinds of thresholds as in the light changing and cars moving example. Otherwise, she considers the example to be “just a bunch of thresholds.”

In this episode, Subset Brainstorming is less productive than Characteristic Oriented brainstorming. I do not have enough data to say whether one or the other type of brainstorming would generally be more or less productive. Both types appear to have potential to help in the brainstorming process even if Group A did not ultimately find a productive example in Subset Brainstorming. Characteristic Oriented Brainstorming does have the strength that any new candidate examples that are brainstormed have at least one of the multiple characteristics one wants. Subset Narrowing has the strength of getting out of the narrower space of examples that have only one particular characteristic in common. This suggests the two types of brainstorming are complimentary.

**Episode Two: Student G Names and Populates The Subcategory of Gradual Thresholds**

I provide data from Student G naming a new subcategory to highlight similar brainstorming practices to Group A. I will first summarize the episode, provide the transcript, and then analyze the data. Although Student G’s initial sense-making of the example is quite interesting, this is provided to help the reader understand the moment of naming a new subcategory and populating it. This chapter’s analysis focus is about the latter phase.
Summary and Transcript

This episode occurs early on in Student G’s workshop and is a little over five minutes in length. She has crafted the definition, “There is a continuous action on an object, a limit is reached where the action has reached a maximum and causes a sudden reaction. This limit would be the threshold.” The first example that the interviewer asks her to consider is “milk going bad in the fridge.” She first rejects the milk example as counting as threshold, re-considers it in light of her definition, and ultimately accepts it and decides to name a new subcategory of “gradual thresholds.” She also tries to populate it with another example (cutting down trees affecting an ecosystem).

Student G will reference the “emu and ostrich” chapter. Recall that the interviewer framed the activity as a bird book with a central definition as well as the potential for chapters about different kinds of birds such as “emus and ostriches.” The transcript follows:

Student G: I wouldn’t think of milk having a threshold. [pause] But, I guess in the sense that there’s a change or a reaction to something, I mean, I guess the fact that it’s just sitting there and is no longer fresh or whatever that means, and then it goes bad. Huh. [pause] So I guess in that case I don’t know what the continuous action would be but the sudden reaction would be that it goes sour. But even that’s not a sudden reaction ‘cuz you can kinda like, ‘oh it’s starting to go bad I’m gonna to toss this’ it’s a gradual
Interviewer: It’s more gradual?
Student G: Yeah
Interviewer: Okay, so, does it seem like something you just wouldn’t want to include if you’re helping people identify threshold stuff?
Student G: I wouldn’t think of milk going sour as having much of a threshold. Maybe it does, but not, it would be a little difficult to grasp at first, maybe it would be in the other chapter.
Interviewer: It’d be in the emu chapter?
Student G: Yeah, the ostrich chapter. [laughing] Yeah. But I wouldn’t know what would the larger overarching category be.
Interviewer: So it’s, you’re a little–
Student G: Gradual processes? ‘cuz I guess there’s other, if we’re thinking about stuff going sour, there’s this gradual process of how many trees you can cut down before it affects the ecosystem and that one, I guess, is also gradual
Interviewer: Right, so it’s not, it’s hard to point to a particular tree or something
Student G: It’s not like, you cut down this tree, and then it all goes down hill
Interviewer: Yeah, it’s kind of more...
Student G: Gradual
Interviewer: A gradual sort of thing. Does it feel like it should be in the emu chapter or not in the book?
Student G: Well, emus are still birds
Interviewer: Right
Student G: And I guess, milk going sour and trees affecting the ecosystem is also a threshold, just not how people typically think of thresholds. And if I were going to ask a kid, “name a bird,” probably not the first bird they’re gonna say is an emu, penguin, or ostrich. But it is a bird and every now and then I’m like ‘oh yeah, penguins are birds’
Interviewer: *Right, they're cute and weird*
Student G: *Yeah, so I guess they should be in the book, but not used to give people a definition and a general understanding of what it is.*

**Analysis: Criteria**
With respect to criteria, we see Student G considering her **Audience** in deciding that the sour milk example would best fit into the emu or ostrich chapter. She notes that, “it would be a little difficult to grasp at first, maybe it would be in the other chapter,” and “I guess they [the two examples] should be in the book, but not used to give people a definition and a general understanding of what it is.” She uses the idea of weird birds as an analogy: “every now and then I’m like ‘oh yeah, penguins are birds.’”

Student G is not as explicit in her dialogue as Student N around the need to brainstorm more examples to name a new subcategory. Her dialogue and actions are suggestive of this, however. She puts a name to the new subcategory and her immediate next step is to say, “‘cuz I guess there’s other [examples]...” She then brainstorms another example. She takes a questioning tone “Gradual processes?” and becomes more confident when she can say “‘cuz I guess there’s other [examples].” This is evidence that Student G also has some type of **(Important) Sub-Categories Have More Than One Member** criterion. Is Student G attempting to understand whether this example is a fluke? If she couldn’t have brainstormed other examples, would she have rejected the subcategory? This is not clear from the data, thus the evidence is not as strong as it was for Group A’s process why exactly she seeks out more examples.

**Analysis: Practices in Brainstorming New candidate examples**
Student G’s method of brainstorming has some resemblances to Group A’s. Group A demonstrated two types of brainstorming: Characteristic Oriented and Subset. In examining Student G’s dialogue, her brainstorming appears most like Group A’s Subset Brainstorming. Explicitly, she says that, “if we’re thinking about stuff going sour, there’s this gradual process of how many trees you can cut down before it affect the ecosystem and that one, I guess, is also gradual.” She appears to be narrowing to consider examples of “stuff going sour.” She uses the word “sour” not in a literal sense of only considering examples involving a sour taste, but in more of a metaphorical sense since she comes up with trees affecting an ecosystem.

**Section Two Conclusion**
In this section we examined Group A naming and populating the subcategory of systems of thresholds and compared this to an episode where Student G named and populated the subcategory of gradual thresholds.
We saw evidence for an **(Important) Sub-Categories Have More Than One Member** criterion in judging whether naming a new subcategory was warranted. There was strong evidence for this criterion in Group A’s work and weaker, but some evidence for, the same criterion in Student G’s work.
Both Group A and Student G engaged in brainstorming more examples to populate the possible new subcategory. Group A spent the longest on this process,
demonstrating two types of brainstorming: Characteristic Oriented and Subset. Student G also brainstormed one example through Subset brainstorming. Lastly, both groups were concerned with their audience and brought to bear different kinds of Audience criteria in deciding to work on the sub-definition. Group A mentioned that they thought kids might wonder about the stoplight example and thus it was worth further consideration. Student G made an analogy to penguins in noting that perhaps the milk example was not the first to give to kids in helping them understand threshold, but it was still a threshold example and could be at the end of the book in a “gradual thresholds” chapter.

**Chapter Four Conclusion**

In this chapter we examined Group A’s process of deciding to name the subcategory of systems of thresholds and populate it with examples. I made cross-case comparisons with other participants to highlight the role of various framework elements. This analysis showed the complicated interaction of various framework elements, highlighted competences in students’ practices and suggested heuristics for definitional refinement more generally.

One connection across the sections is participants’ competences with examples. In section one, participants demonstrated an ability to brainstorm multiple specific versions of the same general example and use criteria in judging these versions. In section two, participants used strategies to support themselves brainstorming new candidate examples that would fit within a newly named subcategory.
Chapter 5 – Conclusion

Overall Results
In this dissertation I presented data and analysis from a set of small group workshops and individual interviews that I designed where undergraduate physical science majors engaged in crafting a definition for threshold. In particular, I:

- Presented an analytic framework for understanding the process of definitional refinement that included elements important to the process, how definitional refinement begins, and two important phases: the pre-phase and the refinement-phase (Chapter Three)
- Named and described a set of criteria that students brought to crafting definitions (Chapter Three)
- Named and described a set of productive practices that students engage in with respect to example use (Chapters Three and Four)
- Differentiated the types of knowledge that students bring to making sense of new candidate examples which suggested some general heuristics for definitional refinement (Chapter Four)

My work has contributed to our understanding of the understudied but important area of students crafting definitions.

Educational Implications and Future Work
The data that I collected was rich with many interesting events that one could examine from a variety of angles. In this section, I wanted to pick out three areas where I could provide some interesting excerpts from my data (some new to this dissertation, some already presented) along with an initial commentary. While I will not get into the detailed level of analysis found in chapters three and four, I will suggest some possible practical educational implications and future work and connect to other STEM education research wherever possible.

My empirical study arose out of my detailed analysis of small group and individual physical science undergraduates working on defining threshold-like phenomena. I will progressively zoom out from threshold-like phenomena in particular and will:

- Position threshold-like phenomena as part of a class of ideas such as oscillation and equilibration that are productive to engage undergraduate physical science students in defining. This can support students' skills in making conceptual and mathematical connections across contexts.
- Discuss the epistemological and affective affordances of engaging undergraduate STEM students in crafting definitions.
- Examine the role of tacit influences on classification generally, the importance of uncovering these influences when crafting definitions, and the potential of certain kinds of examples to support people in this work.
Each of these bulleted points will receive its own focused section, though there are clear connections. For instance, epistemology is connected to supporting students in making connections across contexts. I will keep each section relatively focused, however, to allow for depth in discussion.

Crafting Definitions As Supporting Students In Making Conceptual and Mathematical Connections Between Different Areas of Physics

Traditional undergraduate physics curriculum is taught through separating out areas such as mechanics, thermodynamics, electricity and magnetism, optics, and quantum mechanics. In this type of curriculum it is rare that undergraduates are given the opportunity to make connections across courses, and yet, doing so can bring them insight into conceptual and mathematical commonalities underlying the modeling of many different phenomena. Crafting definitions for the purpose of classification, a process that necessitates the consideration of a diverse example space, holds much promise in supporting students in making these connections.

Threshold-like phenomena are part of a class of ideas such as oscillation and equilibration\(^{12}\) that can be seen again and again across the various subject areas taught in undergraduate physics. Indeed, Feynman, in his discussion of harmonic oscillators in his Lectures, makes note of this very fact (Feynman et al., 1977):

“In the study of physics, usually the course is divided into a series of subjects, such as mechanics, electricity, optics, etc., and one studies one subject after the other. For example, this course has so far dealt mostly with mechanics. But a strange thing occurs again and again: the equations which appear in different fields of physics, and even in other sciences, are often almost exactly the same... so the study of a phenomenon in one field may permit an extension of our knowledge in another field. It is best to realize from the first that such extensions are possible, for otherwise one might not understand the reason for spending a great deal of time and energy on what appears to be only a small part of mechanics.

The harmonic oscillator, which we are about to study, has close analogs in many other fields; although we start with a mechanical example of a weight on a spring, or a pendulum with a small swing, or certain other mechanical devices, we are really studying a certain differential equation. This equation appears again and again in physics and in other sciences, and in fact it is a part of so many phenomena that its close study is well worth our while. Some of the phenomena involving this equation are the oscillations of a mass on a spring; the oscillations of charge flowing back and forth in an electrical circuit; the vibrations of a tuning fork which is generating sound waves; the analogous vibrations of the electrons in an atom, which generate light waves; the equations for the operation of a servosystem, such as a thermostat trying to adjust a temperature; complicated interactions in chemical reactions; the

\(^{12}\) These “patterns” were originally put forth as a way for middle and high school students to study physics by Andrea diSessa and his Patterns Research Group at the University of California, Berkeley.
growth of a colony of bacteria in interaction with the food supply and the 
poisons the bacteria produce; foxes eating rabbits eating grass, and so on; all 
these phenomena follow equations which are very similar to one another, and 
this is the reason why we study the mechanical oscillator in such detail."

Telling students that the mechanical oscillator is worth a great deal of study because the same math underlies so many systems is an important first step. However, it is also important to support students to make these connections themselves which is an important scientific skill. My workshop on threshold-like phenomena provides a template for supporting students to make connections across diverse contexts.

I will now examine some of the conceptual connections my participants made. By design, I started the students out in a somewhat diverse example space that included water boiling as well as blocks falling over the edge of tables and pencils breaking under stress. Over the course of two hours, however, the students far exceeded my expectations in making connections. They brainstormed and explored a set of incredibly diverse examples: wolf and sheep predation, ecosystem damage, water freezing, the body’s response of “getting full” from food, sending rockets and satellites into space, chemical precipitates, blocks overcoming friction, and many more. They connected across thermodynamics and mechanics as well as chemical reactions and biological systems. They noted relevant and subtle distinctions such as reversibility and irreversibility, that some state change happens instantaneously while other state change happens over longer timescales, and that sometimes it is easier to notice a build-up while in other cases it can be imperceptible or missing altogether.

My groups did not spend as much time exploring the underlying mathematics involved in threshold during the design task, but I have some data that suggests that one additional prompt could support them to do so. My framing of the design task as “helping younger people identify threshold in the world” was productive in that it supported my workshop participants in considering a wide variety of examples that people might encounter out in the world. They also uncovered a lot of their own intuition about their perception of examples that they may not have been uncovered had they quickly jumped into the mathematics. Since “younger people” may not have the requisite mathematical understanding to interpret things like graphs and equations, my participants did not explore the underlying mathematics in detail. I had a follow-up conversation with Group A after they had finished the design activity where I asked them about how the definition might change if it was for more physics-experienced audiences. They mentioned that they would no longer have to stay away from language like “state change” and then explored some interesting mathematical intuition where they connected across mechanics and thermodynamics. I will provide a brief excerpt from this data and then discuss how prompting students to consider other audiences has the potential to support them in considering underlying mathematical intuition.

In this excerpt, Student N will explain her intuition about graphs and equations with an example of a block being pushed off of a table. Her group mates will then make connections to water changing state. This dialogue occurs after the interviewer asks the group about a more physics-experienced audience for their
definition. The group mentions that state change was how they had originally discussed the idea before changing to language for younger people. Student N continues discussing state change:

Student N: Change of state is, like, what we said. When two things equal each other, two states are the same, so like, two lines crossing each other on a graph [both Student N and Student L bring their arms up and cross them in the air] sort of is the thing I had in my head. ‘Cuz that’s how you would solve it in real life if you were to be like, ‘when is this pencil gonna fall [as I push it off of the table]?’ you’d be, like, okay, that point in time [Student N taps her pointer finger on the table]

The facilitator then asks Student N to clarify what she means: “Can you draw...?” Student N then draws the graph in Figure 17, below, and narrates: “If you have this equation [draws the line that starts in the top left corner and is fairly straight] and this equation [draws the line that begins in the bottom left corner and whose slope is slightly more parabolic] this is the point [points to intersection] where something could go from doing that to doing that, I guess. That’s just what I was thinking in my head, when we were thinking of changes of state, but like I dunno, this may not actually be how you solve it.”

![Figure 17](image.png)

Figure 17. Student N's graph drawn in response to the Facilitator’s request for further explanation. The arrow is drawn during a later discussion and not as part of the transcript, above.

Upon seeing the graph, Student M makes a connection to thermodynamics and notes, “It's kinda like the water state graph.” Her intonation suggests that she has just noticed this connection and that it is interesting to her. Student L agrees and adds, “Yeah, the boiling points are like that, right?” The group then engages in a discussion about what parameters should be graphed in mechanics and thermodynamics contexts to be able to calculate threshold points.

In my experience teaching, certain graphs in thermodynamics (e.g. the triple point of water) are notoriously hard for undergraduates to interpret and use properly. I’ve seen this over the years as a graduate student instructor of both lower and upper-division undergraduate thermodynamics and statistical mechanics. I will not get into the details here, but student difficulties with and resources for understanding representations is an active area in STEM education. Group A's
dialogue suggests that students might be able to leverage knowledge about critical points in mechanics contexts to better understand what is happening in graphs in thermodynamics. This would be an open question for further study.

I have provided evidence that it is possible for undergraduate physical science students to make important connections both conceptually and mathematically across many diverse examples of threshold-like phenomena within two hours. Those in charge of physics undergraduate curricular decisions have, of course, many competing demands on class time. However, it is worth considering the inclusion of activities where students engage in crafting definitions for ideas like threshold, oscillation, or equilibration, even if it is only possible to do so once or twice during an undergraduate curriculum. Doing so can support students in developing a skill essential to the practice of science: making conceptual and mathematical connections across contexts.

**Epistemological and Affective Affordances of Undergraduates’ Crafting Definitions**

In this dissertation I have drawn a distinction between definitions whose purpose is classification and definitions whose purpose is modeling and experimental measurement. In this section, I want to highlight a commonality: engaging students in crafting definitions, whatever their purpose, has particularly important epistemological and affective affordances. I will provide some excerpts from my own data in making this point. I will also rely on research by Atkins, Salter, and Frank (Atkins and Salter, 2010; Frank and Atkins, 2013) to make this point in the context of definitional activities whose purpose is modeling and experimental measurement.

Epistemology, learning approaches, and affect are interconnected. While not focused on definitions in particular, there is some research showing that bringing students into more authentic scientific practices can support the development of productive attitudes about how to learn physics (Brewe et al., 2009). Authentic science connects to affect as well in that it is often interesting and hard and, if one is properly supported, incredibly engaging. Engagement overlaps with, but is not the same as, “fun”: like going on a long and exhausting hike, there might be fun moments, but it is often about the journey, the development of new skills, the discoveries, and the view at the end.

While affect, epistemology, and approaches to learning were not my explicit research focus, one question that I asked at the end of my interviews sheds some light on these topics and suggests directions for future work. I asked my participants whether the activity they had just participated in would be good to do with freshman. Both Student G’s and Student Q’s response to this question, each enthusiastic that it would be good, show that they wanted more out of their own introductory physics experiences.

I will first provide excerpts from Student G and Student Q’s responses without commenting on affect, epistemology, or approaches to learning. Then, I will break out each into their own sections, comment on the excerpts, and connect to other research on students crafting definitions. Recall that Student G and Q were participants in an educational program that took an approach that supported metacognitive skill building, which likely influenced their ability to articulate these ideas.
**Excerpt One: Student G**

Student G’s response about whether to do this type of activity with freshman was “Yes!” She then discussed how crafting a definition for something like *threshold* could support students in learning effective communication skills, connecting to real-world ideas, and getting a chance to explore and play around with physics, something she did not get out of her introductory physics course:

“Yeah, I think it’s a good idea for them to play around with something they’ve had direct experience with, and effectively communicate their ideas, maybe some people have this intuition about physics and about how things work or why things work, but when they have to explain it to someone they’re kinda at a loss for words. So I think dealing with something that's as simple or complex as threshold or something that people have direct access to and experience with, having to verbalize their thought process and having to form a definition on their own is really powerful in helping people become more effective communicators…

[I also] think it’s kinda important to engage with material other than just doing problem sets and examples which I think in the [introductory calculus based physics course] it’s basically just lecture and problem sets and there’s not really, um, there’s not much room for exploration. Which I understand is to a certain degree okay ‘cuz there’s a certain number of, like, things they need to teach us…but I would definitely like to have some, like anything, optional: ‘Hey come on Saturday! We can come play around with cool physics stuff’”

**Excerpt Two: Student Q**

Student Q also expressed frustration about how college physics is taught in a different way: that her tendency to question “why” was often rebuffed. She then discussed that the definitional activity she participated in was really interesting since she could think more about “*what it means for something to be something or not something.*” Her response to the question about doing the activity with freshman was:

Student Q: Yeah! You always kinda get, physics is, like, presented to you, like, this is how it is and you’re never really encouraged to think about why certain distinctions are made and stuff…[two minutes of dialogue giving a particular example of learning about kinetic versus static friction]…You’re just kinda told ‘yeah, it’s different,’ but you’re like ‘why?’ and they’re like ‘cuz it is.’

Interviewer: So this type of thing might serve to engage students?
Student Q: Yeah, it was really interesting!
Interviewer: Cool!
Student Q: You don’t really get to think about what it means for something to be something or not something
Affect: A Focus On Engagement

Affect is a complex construct and I will pull out the strand of engagement in particular. In the excerpts above, we see Student G and Student Q discuss how engaging the definitional activity was. Student Q says, “It was really interesting!” and Student G mentions that she would come in on a Saturday for it. Although I did not ask the follow-up question about freshman to my group workshops, their engagement could be seen in other ways such as their careful attention to detail, even two hours into working, their frequent joking and laughter together, and, weeks later, members of Group A mentioning that they would enjoy helping me facilitate this type of activity if I did it again. A participant from one of my pilot studies also remarked to me a few weeks later, “I see threshold everywhere now!”

This last pilot student quote connects to aspects of Pugh’s construct of a “transformative experience” or TE. He explains that a TE, “may be defined by three principle qualities: 1) active use of the concept, 2) an expansion of perception, and 3) an expansion of value. Active use means the individual seeks out or takes advantage of opportunities to use the concept as a potential lens for more fully perceiving the world” (Pugh, 2002). This construct broadens engagement outside of students’ immediate classroom experience and looks for evidence of students using ideas learned in the classroom out in the world. Although my research focus did not involve systematically following up with my own participants about the extent to which the workshop acted as a transformative experience, Frank and Atkins (Frank and Atkins, 2013) have evidence that students in their college science courses experience aspects of having transformative experiences. Students in Atkins’ inquiry science course for pre-service elementary school teachers, in particular, "had the highest percentage of agreement” on TE survey items, which was 89% agreement. While it is not possible to make a direct link between definitional activities and TE in their courses since TE surveys asks general questions about the course overall, the crafting of definitions was a central part of the authentic science that students engaged in in the inquiry course and was likely a contributor to students’ high TE survey scores.

Atkins and Frank’s work, as well as my own, gives some preliminary evidence of the potential of certain types of definitional activities to support student engagement both within and outside of the classroom. Future work might further examine engagement in connection with developing important scientific skills and supporting more productive beliefs about what physics is and how to learn it.

Epistemology And Approaches To Learning

In this section I want to focus on two aspects of epistemology and approaches to learning. First, I want to examine how STEM students often memorize or even reject definitions that they do not understand rather than having a good understanding of their use. Second, I make the point that most traditional introductory physics courses negatively affect epistemology in the sense that students believe physics has less to do with the real world and that it is not important to make sense of their own intuition. There is some preliminary evidence that my workshops can support more realistic and productive beliefs about definitions as well as encourage students to see benefits to sense-making.
In the context of scientific inquiry classrooms, Atkins and Salter make the argument that “in a course where the goal is to engage students in doing science,” it is “more authentic” to “allow for theoretical and empirical progress before converging on a precise, consensus definition” (Atkins and Salter, 2010). Bringing students into crafting definitions can support them in coming to understand that definitions are not perfect things to be memorized, rather, they are crafted by humans who must negotiate many sometimes conflicting design criteria, and are more or less useful toward particular goals and contexts. Yet, even in many reformed introductory college physics curricula such as the University of Washington Tutorials or Modeling Instruction (McDermott and Shaffer, 2002; Brewe, 2008), it is rare that students spend time crafting definitions for themselves.

If we never engage students in crafting definitions, they may not understand their use and ultimately reject or forget them. Although it is in the context of a younger group of grade school math students, a brief excerpt from a paper by Nachleili and Sfard makes this point well (Nachleili and Sfard, 2003). In this geometry excerpt, reproduced in Table 3, below, a teacher is having his or her students classify examples of kites. One student, Noam, can recite the definition of kite and understands that a particular geometrical figure fulfills it, however, “he refuses to accept the definition as the ultimate touchstone for determining whether the given case belongs to the category of kites.” The dialogue from their paper follows:

<table>
<thead>
<tr>
<th>WHAT IS DONE</th>
<th>What is said</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points to that figure:</td>
<td>[1] Teacher: Is this figure a kite?</td>
</tr>
<tr>
<td></td>
<td>[2] Noam: No</td>
</tr>
<tr>
<td>Points to that figure:</td>
<td>[6] Noam: like this</td>
</tr>
<tr>
<td></td>
<td>[7] Teacher: What is the definition of a kite?</td>
</tr>
<tr>
<td></td>
<td>[8] Noam: Reads the definition. According to the definition this is a kite but I know it is not.</td>
</tr>
</tbody>
</table>

Table 3. A table of transcript from a geometry lesson reproduced from Nachlieli and Sfard’s work on the activity of defining or AoD (Nachlieli and Sfard, 2003).

My interpretation of this episode is that Noam makes some implicit judgment that communication outside of the mathematics classroom is more important to defining kite when he states, “it doesn’t look like one.” He could be cognizant of why he makes this judgment, though there is not enough data to say. Either he does not understand the mathematical context in which his class is using the definition for kite, or he has not been convinced of its importance. Similarly, many college physics
students can memorize and recite definitions but promptly leave them behind once the class is over (Redish, 2003).

My workshops were different: students employed their crafted definition in making sense of new candidate examples and, by the end of the workshop, judged their definitions positively. In particular, I asked them to fill out a final survey where they judged their definitions on a scale of “everything I think should be included in the definition is there” to “this definition isn’t at all reasonable” (See Figure 19, below). All students ranked their definitions at least at the mid-point between “neutral” and “Everything is there!” and some even circled “Everything is there!” I also asked the survey question, “Should anything be added to the definition?” Two students noted the difficulty of including the idea that some state change happens more gradually rather than instantaneously. Some of the responses are below:

- “No I think it is a good definition, it is broad and encompasses a lot.”
- “example of thresholds that take a long time”
- “I would like to include something about the gradual process. How thresholds can be reached gradually though I am unsure on how I would write it in a coherent and concise manner.”

(1) Please rate the final definition on a scale of “Everything I think should be included in the definition is there” to “this definition isn’t at all reasonable.”

Not reasonable!<-----------------------------Neutral------------------------> Everything is there!

(2) Should anything else be added to the definition?

I would like to include something about the gradual process. How thresholds can be reached gradually though I am unsure on how I would write it in a coherent and concise manner.

Figure 19. Student G ranked her definition highly but also noted some additional changes she could work on to make it better.

Student Q and G’s enthusiasm and arguments for bringing the threshold defining activity to freshman connected to important epistemological issues not just about definitions in particular but about sense-making generally. One hopes that students move away from the idea that learning science “consist[s] mainly of absorbing information” and come to see it as, “rely[ing] crucially on constructing one’s own understanding by working through the material actively, by relating new material to prior experiences, intuitions, and knowledge, and by reflecting upon and monitoring one’s understanding” (Elby, 1998). There is much evidence that traditional introductory physics courses negatively affect students’ classroom attitudes around what is important to learning science and can even change their approaches to learning in negative ways (Hammer, 1989; Hammer, 1994; Redish et al., 1998; Adams et al., 2005). However, Student Q noted the importance of sense-
making when she said that the threshold activity allowed for something she had not found in her introductory courses, that she got to, “think about why certain distinctions are made.” Student G highlighted the importance of sense-making activities connected with students’ direct experiences of the world: “I think it’s a good idea for them to play around with something they’ve had direct experience with...maybe some people have this intuition about physics...but when they have to explain it to someone they’re kinda at a loss for words.”

Student Q and G’s quotes give some preliminary evidence that activities like my threshold workshop might support students in developing productive attitudes about how to learn physics with respect to seeing the importance of sense-making and drawing on their real world experiences. I also made the argument that engaging students in crafting definitions, where they must name and negotiate criteria for what makes a good definition, has the potential to support them in developing more realistic and productive beliefs about definitions and their role in science and science learning. They are also more likely to use a definition in practice.

**Future Research Direction: The Importance of Engaging in Crafting Definitions To Uncover Tacit Influences On Classification And The Role of Particular Kinds of Examples**

In this section I want to first note the importance of uncovering tacit influences on classification when one is crafting a definition. I will then discuss some heuristics for accomplishing this, first in the context of law education and then with some specifics from my own data. I will show that particular kinds of examples help in this work. While the data and commentary in this section bears some resemblance to my core research as I highlight some of my framework elements, I want to stress that these are initial ideas. This is the future research direction that is most closely tied to the core analysis I presented in this dissertation.

As I explored in detail in my introduction, we know from cognitive science, linguistics, and physics education research that our ideas and intuitions are affected by typicality and framing effects from our exposure to and interaction with the physical and social worlds (Rips, 1975; Lakoff, 1987; diSessa, 1991; Minsky, 1975; Tannen, 1993). Since crafting a productive definition for classification involves uncovering these tacit influences and deciding whether and how they should be incorporated, how does one go about this?

Research on engaging law students in crafting law definitions suggests that an interplay between the crafted definition, a base case, and set of “hypotheticals” where only one aspect of the base case is changed at a time (e.g. the location) is important to uncovering tacit influences on classification (Rissland, 1983). I saw a similar interplay in my own data. Although my workshops did not involve base cases and hypotheticals, particular examples that had only one aspect that participants might feel was “off” worked in a similar way.

For threshold-like phenomena, two timescales often affected whether an example felt “off” to participants and thus were important to uncover. These two timescales were the length of the build-up phase and the length of the change or reaction. If an example had a build-up that happened imperceptibly fast or if the change did not happen instantaneously, participants often expressed discomfort at
its fit. Examples with two non-prototypical timescales were more difficult for participants to make sense of than examples with only one non-prototypical timescale.

I will now return to the first few lines of an episode of definitional refinement that was explored in Chapter Three. I will re-examine it as an instance of a student uncovering a tacit influence on her classification through an example with only one aspect that felt “off.” In this episode, Student Q will become aware that her perception of build up is affecting her classification process with help from the example of sneezing. At the beginning of this episode, she had a crafted definition that did not include the idea of build up: “A finite point where something turns into something else, or the point at which something happens when you act on an object.” After she becomes aware of the importance of build up, she refines her definition to include an additional phrase about it: “Two ways of seeing threshold: some point that is built up to, or just a barrier.”

In the transcript excerpt, below, the interviewer had just asked Student Q about the example of sneezing and her response follows:

Student Q: There’s a point at which you have to sneeze and, let’s say it’s more like coughing, you can take a little bit of annoyance in your throat to the point where you’re like coughing up organs and stuff and there’s kind of like a trigger point almost, whereas, I guess there is, too, for sneezing, but, I dunno, my sneezes always catch me by surprise usually.
Interviewer: Yeah so it seems, so sneezing seems a little less
Student Q: There’s no build up, almost, well, probably there is, but not that I’ve ever noticed.
Interviewer: yeah
Student Q: Like, you put strain on a pencil and it can take some, but there’s a point at which you can’t [pause] sneezing seems to just come out of the blue. Maybe my sneezes are weird, I don’t really know.

In this episode, we see Student Q able to articulate what feels “off” about sneezing: “there’s no build up...” Sneezing is a good example for Student Q in uncovering this tacit influence since her personal experience of sneezing does not involve a build up. For other students, their personal experience might involve more of a build up and the example might not support them in uncovering the potential influence of this timescale on classification.

This is an impressive episode overall: Student Q not only articulates what feels “off” about sneezing, but she refines her definition to highlight that it is not always possible to perceive a build up. She eventually\(^\text{13}\) adds an additional phrase to her definition: “Two ways of seeing threshold: some point that is built up to, or just a barrier.”

\(^{13}\) The rest of this episode, including the details of definitional refinement, can be found in Chapter Three.
Student Q’s episode gives us one way that a student can uncover tacit influences on classification: by noticing Outside DUK\textsuperscript{14} use in making sense of an example. The sneezing example was also special: it had only one aspect that “felt off,” which supported Student Q in articulating that her perception of build up affected her classification of examples.

An additional way that students can uncover tacit influences is through recognizing that a part of their definition might be unnecessary in considering an example that feels “off.” One of my students, Student G, had used the phrase “\textit{sudden reaction}” in her definition. This use of the adjective “\textit{sudden}” is connected to the human perception of timescale of the change or reaction in a threshold event. Student G called into question the necessity of the “\textit{sudden}” adjective through interacting with a particular example: milk going bad in the fridge.\textsuperscript{15} Unfortunately, Student G also thought that the example also had an unclear build up mechanism, leading to some confusion between the two timescales. Although the example suffered from the complication of multiple aspects feeling “off,” her episode suggests an additional way to uncover tacit influences. If an example meets all of a crafted definition except for something like an adjective or adverb, it is worth considering whether the adjective or adverb is connected to human perception. In this case, “\textit{sudden}” was representative of a human perception of timescale.

The most useful examples to support students in uncovering tacit influences are ones where only one aspect feels “off” to them. Therefore, if one wants to facilitate activities where students craft a definition, supporting them to consider these special examples is critical. This might mean supporting the group to brainstorm a diverse example space so that they are more likely to come across these types of examples. Sometimes the facilitator might have some good guesses, though, and can introduce them into the activity. After I facilitated a few pilot workshops, I came across examples that often worked to help students uncover tacit influences (e.g. sneezing and milk going bad). However, it is important to note that individuals can have very different personal experiences with particular examples.\textsuperscript{16} While sneezing worked well for Student Q to uncover the idea of build up, other students might have an experience of sneezing that involves a build up. This means that examples cannot be assumed to work for everyone.

In these episodes we have seen the critical role of particular kinds of examples in participants’ ability to become aware of tacit influences on classification. This awareness can support them to change their definition in important ways.

\textbf{Conclusion}

In this last chapter I summarized my research findings and then explored some possible educational implications and future research directions. I explored

\textsuperscript{14} Recall that Outside DUK is definitionally unarticulated knowledge with no relationship to any statements in the current form of the definition. See Chapter Four for a more detailed description of various types of DUK.

\textsuperscript{15} More details from this particular episode can be seen in the second half of Chapter Four.

\textsuperscript{16} These differences in personal experiences with examples were explored in more detail in Chapter Four.
educational implications for college physics teaching as well as more general instances where one is interested in crafting a definition for the purpose of classification.
References


Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of verbal...
learning and verbal behavior, 14(6), 665-681
Swanson, H. (In preparation 2013). Instruction Designed to Support the Development of Students’ Knowledge of General Patterns that Apply Across Domains.