# EXPLORING THE COMPLEX RELATIONSHIPS AMONG REASONING, CONTENT UNDERSTANDING, AND INTUITION IN PHYSICS

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Cody Ray Gette

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Physics

February 2019

Fargo, North Dakota

# NORTH DAKOTA STATE UNIVERSITY

Graduate School

**Title**

EXPLORING THE COMPLEX RELATIONSHIPS AMONG REASONING, CONTENT UNDERSTANDING, AND INTUITION IN PHYSICS

**By**

Cody Ray Gette

The supervisory committee certifies that this dissertation complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Prof. Mila Kryjevskaia
Chair

Prof. Sylvio May

Prof. Warren Christensen

Prof. James Nyachwaya

Approved:

02/25/2019
Date

Prof. Sylvio May
Department Chair

## ABSTRACT

Physics education research over the past few decades has made significant advances toward improving instructional practices and developing effective instructional materials for physics classrooms. In some contexts, however, after multiple instructional refinements difficulties can remain persistent. Recent findings in PER suggest that many of these difficulties are consistent with reasoning paths accounted for by dual process theories of reasoning. Students often appear to be able to employ correct conceptual understanding in one context, but neglect to demonstrate the same understanding in another closely related context. This thesis explores the use of dual-process theories of reasoning as a lens for interpreting observed patterns of student reasoning. First, we examined both the impact of problem design and the impact of instruction targeting accessible intuitive ideas in the context of sinking and floating. We found that targeted instruction which directly addressed everyday experiences had a significant impact on student performance in that context. We also found that changes to problem design and instruction emphasizing correct approaches had little impact on performance. Further investigations into the cognitive mechanisms behind student reasoning patterns found a positive relationship between student cognitive reflection skills and performance on Newton's third law problems. Findings suggest that those with higher cognitive reflection skills, as measured by the CRT, are more likely to 1) answer correctly on problems which elicit intuitively appealing but incorrect answers 2) provide correct and complete physical justification to problem solutions and 3) answer problems consistently. Finally, we examined student reasoning patterns in the context of mechanical waves. We attempted to influence intuitive approaches with video simulations of the physical situation. We found that students tended to reason with mathematical approaches and had difficulty overcoming intuitive ideas even after viewing the physical simulation.

# ACKNOWLEDGEMENTS

I would like to thank my academic advisor, Dr. Mila Kryjevskaia for her support throughout my academic career. From my time as an undergrad at NDSU, she has always supported my ambitions and provided insight and guidance wherever my next step has taken me. Her patience and willingness to always take time for discussions and feedback has been invaluable to me as a graduate student.

I would like to express additional gratitude to Dr. MacKenzie Stetzer and Dr. Andrew Boudreaux for providing considerable discussions and feedback on nearly every aspect of my work including publications and presentations. And to Dr. Beth Lindsey and Dr. Paula Heron for their contributions and feedback on the instructional designs we implemented as well as feedback and valuable discussions.

I would like to thank my committee members: Dr. Sylvio May, Dr. Warren Christensen, and Dr. James Nyachwaya for their comments and critiques of my research.

I would like to especially thank Catt Miller for her unwavering confidence, countless discussions, and for always keeping me grounded. I would like to thank Kyle Strand for innumerous cups of coffee and support on and off campus, and Sumeyye Suri for her continuous encouragement and correspondence from overseas. I would also like to thank my sister, Megan, for always inspiring me to pursue what I find meaningful. Finally, I would like to thank my parents for always being willing to listen and provide support while we sit and watch the waves. Without them, none of this would have been possible.

# DEDICATION

*This thesis is dedicated to my mom, Gayle, and my dad, Mel.*

# TABLE OF CONTENTS

# LIST OF TABLES

Table Page

# LIST OF FIGURES

# LIST OF APPENDIX FIGURES

Figure | Page

# 1. INTRODUCTION

Over the past several decades, physics education researchers have identified student difficulties with many physics concepts and developed instructional strategies that significantly improve student performance in many contexts [3–9]. Analysis of incorrect responses often revealed patterns that may be interpreted as stemming from the students' lack of conceptual understanding (*i.e.,* mindware). At the same time, researchers identified many contexts in which multiple refinements of instructional materials designed to address conceptual difficulties did not necessarily lead to significant improvements in student performance [10–13]. A number of recent investigations have revealed that some of these patterns of persistent incorrect responses may be due to reasoning difficulties rather than to a lack of relevant conceptual understanding [10, 11, 14–20]. In order to help students develop more productive thinking habits in the context of physics (and beyond), it is imperative to direct efforts toward (1) pinpointing more precisely the underlying mechanisms that may lead to observed patterns of student reasoning, both productive and unproductive, and (2) utilizing the results of this research in order to develop instructional materials designed to enhance student reasoning skills.

This thesis explores the relationships between student reasoning tendencies, content understanding and intuitive ideas in physics classrooms. We draw from research in cognitive psychology to better understand the reasoning mechanisms that students appear to use in solving physics problems. The theoretical ideas from Dual-Process Theories of Reasoning (DPToR), accessibility, cognitive reflection, and mindware are used as a lens for interpreting student reasoning processes. We examine the impact of problem design, targeted instruction, cognitive reflection skills as well as mathematical intuition. The results point to cognitive tendencies and skills in physics classrooms consistent with DPToR as well as represent a first step in developing strategies for utilizing theories from cognitive science as a guide for refining instructional materials. Presented below are the theoretical ideas that play an important role in our understanding of student reasoning patterns as well as the current physics education literature that serve as a basis for the investigations.

### 1.1. Decision making

Everybody makes decisions. Generally speaking every aspect of everyday life is influenced by a series of small decisions. Under normal circumstances, our everyday lives are relatively unaffected if we have a small, occasional "poor" decision. What qualifies as a "poor" decision is subjective in many contexts, but in general we are able to successfully cross the street, recognize friends, and perform simple calculations when the situation requires it. Our decision making is effective and generally we proceed with confidence, however we do not always proceed rationally.

Rational thought can be broadly defined as conscious reasoning in logical steps, without bias. Cognitive scientists studying heuristics and biases have identified many systematic tendencies of normally rational people acting in irrational and biased ways. Evidence from this research has shown that rather than acting as rational beings, people tend to base decisions on many intrinsic biases and intuitive thoughts which change from context to context. For example, if a subject is asked a question and imagines themselves in a situation that is beneficial to them they will tend to answer differently than if they imagine themselves in a situation that is decidedly unfavorable to them, but with the same outcome [21, 22]. In addition, people tend to look for support to answers they already believe to be true [1, 23, 24] (See also the Wason Rule Discovery Test [25]). Probabilities are particularly difficult for people to judge and responses are affected profoundly by information provided just before the question is presented, regardless of its relevance [24, 26, 27].

Much of the research in heuristics and biases can be succinctly summarized as "generally rational people thinking irrationally." Even people who are very intelligent by a number of intelligence measures can find themselves *acting* in irrational ways [28]. The question then becomes not "why do people make poor decisions?," but rather "under what circumstances do people who are capable of making good decisions, make poor ones instead?."

This question is of especial interest to those who depend on their decision making abilities in a high stakes environment. Stock brokers, businesspeople, and students taking a timed exam could particularly benefit from good decision making and reasoning. It would be beneficial to all fields which encounter reasoning difficulties to explore what mechanisms are behind biased responses, and how they affect our abilities to apply our knowledge to a given situation. Many current theories of reasoning developed in cognitive psychology propose that cognition relies on a dual-process system [1, 29]. The interplay between the two processes generates reasoning chains which draw on knowledge and experience when attempting to solve a given problem. Whether this problem is deciding when it is safe to cross the street or recognizing the applicability of the principle of conservation of energy, each problem requires a strategy to solve. Understanding when people productively use their knowledge and when they do not is a major part of cognitive psychology and heuristics and biases research. Inconsistent and irrational reasoning is often encountered in the classroom, and research investigating the relationship between current theories of reasoning and performance in classrooms may have important implications for instruction. Understanding the mechanisms behind reasoning approaches will help us to better understand the difficulties students face when learning new content, and consequently design more effective instruction.

### 1.2. Dysrationalia

There are many examples in heuristics and biases research which illustrate one's tendency to not answer rationally, or to use all available knowledge in making a rational decision. A simple example which is often used is the following problem: A bat and ball together cost \$1.10. If the bat costs \$1 more than the ball, how much does the ball cost? Many people, including students at Harvard and MIT, answer 10 cents [28], however this is incorrect (the correct answer is 5 cents). In principle, this problem is straightforward and working it out on paper takes no more than a line or two of math. It cannot be argued that the majority of people who approach this problem are incapable of performing the calculations, so instead an incorrect answer suggests a failure in mental processing. This observation of capable reasoners making seemingly irrational decisions has been given the name 'dysrationalia' [30]. First coined by Keith Stanovich, dysrationalia is defined as the inability to think and behave rationally despite having adequate intelligence. Understanding the causes of irrational thinking and unproductive problem solving can be vitally important to instruction in the classroom. Research into decision making can help guide our investigations into student thinking and ultimately guide the creation of better instruction.

In the book, *What Intelligence Tests Miss*, Stanovich proposes two causes of irrational thinking (1) a processing problem and (2) a content knowledge problem. He suggests that despite our ability to reason effectively with slow, careful mechanisms, we are *cognitive misers*; we choose whatever mechanism will require the least computational effort. A processing problem in many cases results from the fact that a reasoner can apply several cognitive mechanisms, but will default to choosing the one that requires the least effort. As in the bat and ball problem, we tend to go with our initial answer which takes little thought: 10 cents. An important realization of this research is that miserly processing is *independent* of intelligence and even those with "superior" computational processing power can be subject to irrational biases [1, 28, 30].

The second proposed cause of dysrationalia is a content knowledge problem. Coined by David Perkins, the term "mindware" refers to all of the rules, data, procedures, strategies, and other cognitive tools that must be retrieved from memory in order to think rationally [30]. A lack or absence of such knowledge is referred to as a "mindware gap". In other words, if a person has not learned a strategy or rule necessary to reason rationally through a problem, it will appear as if they are thinking irrationally. While there are many factors that contribute to these two causes of irrational thinking, analysis of processing problems and mindware gaps in reasoning approaches can account for many problem solving patterns seen in cognitive psychology.

### 1.3. Bounded rationality

Investigations into the causes of unproductive reasoning approaches is not new. Work done by Simon in 1955 explained reasoning patterns related to judgment and choice in terms of *bounded rationality*, or in other words, that people make decisions that are informed by real-world constraints such as limited time and limited information. This led to the idea that people instead employ fast and efficient heuristics to make decisions. As outlined above, our decision making is generally effective, but we do not always proceed rationally. It is instead thought that we form an initial mental model based on our experiences and what has 'worked' for us in the past.

The initial mental model formed by a problem solver must come when the solver has formed at least some preliminary idea of "what the problem is all about" or what it means to have "solved" the problem. Once this initial state has been established, the solver can then focus on strategies and possibilities to reach the *goal*. Stellan Olsson in the book *Deep Learning: How the Mind Overrides Experience* describes how visual inspection and interpretation of the given information set up an initial problem representation based on the perceptions and prior experiences of the individual [31].

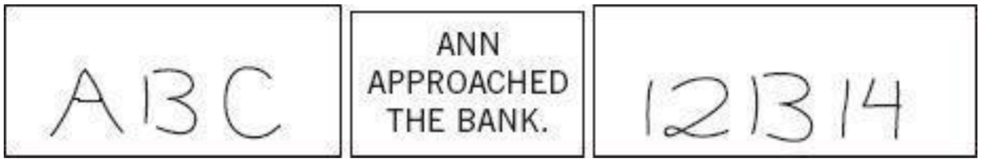



Figure 1.1. Visual representations, taken from Kahneman "Thinking Fast and Slow" [1].

> Whether in the psychologist's laboratory, in school or in work situations, people are confronted with situations and problem materials that require visual inspection and interpretation [...] For a situation or a set of materials to constitute a problem, it must be paired with a goal, a mental representation of a desired (but, by definition, not yet realized) state of affairs. If the goal is posed and communicated by someone else, the constructive nature of comprehension has similar consequences as the constructive nature of vision: Goal comprehension is not uniquely determined by the verbal input. The problem solver's representation of that goal is an interpretation based on prior experience.

From our perceptions, we interpret the problem and form an assumed goal. To experience the speed and ease with which we can form an idea or representation given the context, consider the two lines of written text in Fig. 1.1. Most people will read the middle character in the left and right boxes automatically as the letter 'B' and the number '13', however they are in fact the same character. While these two boxes do not represent a 'problem space' it can illustrate how quickly and subconsciously one can form a representation of the information without considering alternatives or possibilities and only considering the context in which they are presented.

This also illustrates the efficiency of our processing mechanisms. Instead of searching every possibility, we tend to assume that we are correct unless we have a good reason to think otherwise. Taken to extremes, this idea is consistent with describing reasoners as 'cognitive misers.' An efficient way to reduce the number of possibilities is to consider only those which have 'worked' in the past. In other words, we confine our search to only a few promising initial mental models. This approach may highlight a difference between experts and novices. As we gain expertise in an area, we are able to more effectively 'choose' our heuristics so that we consider only those ideas which may be beneficial to making the decision. In fact, experts in a particular content area seem to be able to quickly and efficiently search for and find the best course of action. However, as the generation of initial mental models are generally governed by subconscious processes, experts can have trouble articulating exactly how they know they are acting correctly [32]. Understanding what heuristics are first available to reasoners and what may cue these heuristics can give important insights into understanding how reasoning paths and mental models take shape.

### 1.4. Dual process theories

It is important to note that an individual's assumed goal doesn't necessarily have to align with the goals of the one who posed the question. In fact, since each individual has different prior experiences and no one person has lived an identical life to another, it would seem studying how people make decisions and solve problems would be a fruitless task. Remarkably, despite our differences, people tend to approach problems in predictable ways (See for example [11]).

Cognitive psychologists have noted that for a variety of tasks there appears to be at least two distinct ways to arrive at a response. The difference in paths to responses can be attributed to the interactions between two cognitive processes (Process 1 and Process 2) [1,2,33]. Process 1 (often referred to as the *heuristic* process) is described as quick, intuitive, effortless, and subconscious; it is responsible for developing a "first impression" mental model of (or a way of thinking about) presented situations. Process 2 (often referred to as the *analytic* process) is slow, conscious, and effortful; it is capable of producing logic-based or rule-based reasoning. The theories which explore the interaction between the two processes are called dual process theories. In articles on the heuristic-analytic theory of reasoning, Evans offers a diagram, shown in Fig. 1.2, that provides a visual aid for tracing and understanding various reasoning paths arising from the interactions (or lack thereof) between the two processes [2].

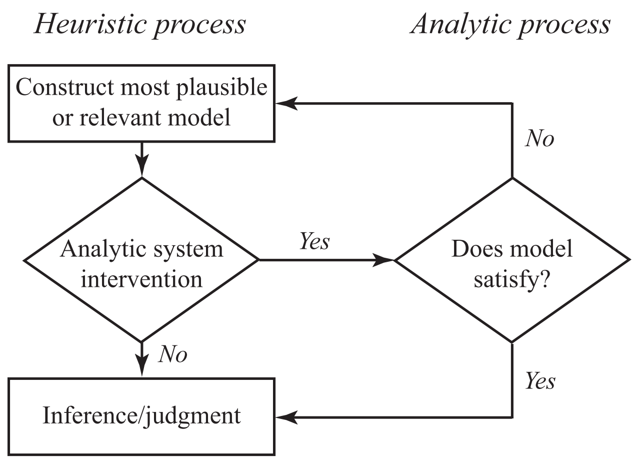



Figure 1.2. Heuristic-analytic theory of reasoning and decision making [2].

Once a reasoner is presented with a specific situation or a task, Process 1 *immediately* and *subconsciously* develops a mental model of the situation based on prior knowledge and experiences, contextual cues, relevance, and other factors. This "first-available" mental model often represents a quick and subconscious attempt on the part of a reasoner to produce a coherent and plausible way of thinking about the situation at hand. In everyday language, it is often referred to as a "gut feeling," while in cognitive psychology the construct of the "first-available mental model" is used to provide a formal, operational definition of a reasoner's intuition. Once the first-available mental model is developed, it becomes available for scrutiny by the more rigorous, analytical Process 2. However, if a reasoner feels confident in the first-available mental model, the analytic process may be circumvented entirely [34, 35]. If so, the first-available mental model yields a final response. This direct thinking path from a "gut feeling" to a final inference or judgment is overwhelmingly prevalent in everyday activities. In general, Process 1 is fairly efficient and accurate at providing a quick assessment of familiar situations. Moreover, even if Process 1 suggests an erroneous conclusion or a flawed action in an everyday domain, we are often not explicitly aware of such failures. Since we do not generally get immediate feedback on our mistakes, it is not surprising that we learn to trust our intuition, which seems to work in everyday life. As a result, novice learners may transfer their reliance on intuitively appealing thoughts into the context of science instruction as well. As such, they may perceive little need for explicit and rigorous validation of their thinking.

In order to catch a mistake, the analytic Process 2 must be engaged and placed on alert. If the analytic process is not satisfied with the current response, the reasoning path returns to the heuristic process, which suggests a new mental model for consideration. It is important to stress that even if the analytic process is engaged, it may be impaired by biases of its own. As such, the engagement of the analytic process does not necessarily ensure that a reasoning flaw will be detected or that a desired, logic-based argument will be generated. For example, people tend to create coherence by actively seeking confirming evidence for an existing first impression idea, a tendency referred to as *confirmation bias* [21]. In contrast to good scientific thinking, reasoners sometimes fail to spontaneously search for alternative mental models or counter-arguments that could potentially falsify their original predictions. Therefore, even when the analytic process has been engaged, a "first impression," intuitive response often still emerges as the final answer.

In summary, according to dual process theories of reasoning, our first assessment of a situation at hand occurs through the eyes of the quick and subconscious heuristic process, which cannot be turned off. We develop "first impression" mental models, often without explicit awareness of thinking processes, and we often construct an argument to support the conclusion suggested by that intuitive model. The intuitive thinking process, which operates based on prior experiences and contextual cues (rather than formal reasoning), creates bias that interferes with the ability of the rational analytic thinking process to construct a valid logical reasoning chain.

### 1.5. Connection with physics education research

Many recent findings in PER are consistent with the reasoning paths accounted for by dual process theories of reasoning. Indeed, it has been found that many students apply correct conceptual knowledge in a selective manner. They appear to be able to employ correct conceptual understanding in order to construct an argument supporting highly plausible and correct conclusions while neglecting to utilize the same correct conceptual understanding in order to refute highly plausible and erroneous conclusions [12, 19, 36]. We hypothesize that some persistent student difficulties identified in physics may not necessarily be due to a lack of conceptual understanding, but rather may stem from basic reasoning difficulties outlined by heuristics and biases research (*e.g.,* confirmation bias).

#### 1.5.1. Fluency and accessibility

Physics education researchers have begun to use knowledge of implicit (*i.e.,* heuristic) processes in order to identify more precisely factors that impact student reasoning in the context of physics [10–12, 18, 36–38]. One such process is related to the *fluency heuristic*, which is linked to processing time and provides the following mechanism for cuing a specific, first-available mental model: the faster an idea is processed, the more "weight" it is given in reasoning [39–41]. Research by Heckler and Scaife on the impact of the fluency heuristic on response patterns is of particular relevance [36]. Heckler and Scaife operationally defined and measured fluency as the time needed to arrive at an answer. They considered a class of tasks that contained competing relevant and irrelevant information. By comparing the time needed to process the given information, they determined that the relative fluency of irrelevant vs. relevant information had a significant impact on answer patterns.

Related to the concept of fluency is the theoretical idea of *accessibility*. Accessibility refers to the extent to which concepts are at the forefront of one's mind and are therefore likely to be used when making judgments [22, 24, 42–44]. Both concepts refer to the ease or difficulty with which information comes to the mind; however, while fluency pertains to the speed of processing (*e.g.,* height might be processed faster than slope), accessibility refers to the likelihood that information comes to mind and is perceived to be relevant (*i.e.,* is used in building an argument). In a recent study, Heckler and Bogdan analyzed student responses to an array of common physical scenarios and provided empirical evidence for the "accessibility rule" [44]. The rule states that "alternative explanations are considered less frequently when an explanatory factor with high accessibility is offered first." In other words, students are more likely to reason with a highly accessible factor, and the higher the accessibility, the less likely they are to consider alternative explanations. They also note that although relevant factors are more likely to be recognized by students than recalled, recognition is not sufficient and "easy recall" (*i.e.,* accessibility) plays a critical role in productive reasoning. In this thesis, we apply the concept of accessibility to the contexts of buoyancy and mechanical waves (See Chapters 2 and 4). We hypothesize that understanding the relative accessibility of student ideas will allow instructors and researchers to more productively predict which ideas are likely to be used by students in the classroom and on exams. In turn, this could influence the refinement and implementation of new instructional materials.

### 1.5.2. Cognitive reflection

In addition to accessibility, we aim to understand the cognitive tendencies of those students who successfully reflect on and overcome intuitive ideas. The mechanistic reasoning paths outlined by DPToR suggest that reasoners who are able to identify their own intuitive thoughts and engage in an effortful, logic-based analytical evaluation of such thoughts are more likely to reason productively.

1. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?
2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half the lake?

Figure 1.3. Cognitive Reflection Test.

The ability to mediate intuitive thinking by reasoning more analytically is often called *cognitive reflection*. As stated above, according to DPToR, the heuristic process operates continuously and cannot be turned off; reasoners interpret the world around them largely through the filter of fast, intuitive processing. Intuitive thinking processes thus play a major role in how students interpret classroom instruction and in how students leverage the conceptual understanding developed as a result of that instruction.

The Cognitive Reflection Test (CRT), presented in Fig. 1.3, has been developed in psychology to gauge a reasoner's tendency to engage the analytic process in order to evaluate the output of the intuitive process [28]. The key feature of each CRT question is that, for most reasoners, an intuitively appealing, but incorrect, answer comes to mind immediately and often without conscious effort. This answer must first be rejected and then replaced with a correct solution. (The correct answers to the three CRT questions are 5 cents, 5 minutes, and 47 days, respectively.) A score of 2 or 3 on the CRT indicates a relatively strong tendency to mediate intuition with conscious reflection. We emphasize that weak performance on the CRT does not generally indicate a deficit in formal reasoning skills per se, but instead suggests a difficulty with activating the necessary reasoning.

The CRT has been validated and widely used in psychology research. Numerous studies have been conducted in order to pinpoint specific aspects of cognition measured by the CRT. Researchers generally agree that the CRT is *not* a measure of cognitive abilities or mathematical skills [30, 45–47]. Toplak et al. probed the relationship between the CRT and performance on a wide array of tasks from the heuristics-and-biases literature selected to "reflect important aspects of rational thought," including probabilistic reasoning, hypothetical thought, theory justification, scientific reasoning, and the tendency to think statistically [48]. Results indicate that the CRT is a more "potent predictor" of performance on heuristics-and-biases tasks than measures of cognitive ability, other thinking dispositions, or executive functioning. The researchers argued that these measures do not gauge the tendency "toward miserly processing in the way that the CRT does."

A growing body of PER literature provides compelling evidence suggesting that the successful reasoning on some types of physics questions, much like a strong CRT performance, requires students to override initial intuitive responses. As such, it is reasonable to predict a positive relationship between the CRT score and student performance on these types of questions. At the same time, the ability to recognize shortcuts it one's own reasoning *alone* is not enough to ensure a strong performance in physics. In the presence of a "mindware gap," even if a student recognizes the intuitive nature of a compelling response, the lack of relevant knowledge is likely to prevent the productive engagement of the analytic process, thus, yielding a poor response. According to Stanovich "Surpressing one response is not helpful unless there is a better response available to substitute for it [49]." Our prior study, aimed at probing the relationships between cognitive reflection and student performance, provides evidence for this claim in the context of introductory physics instruction (See Chapter 3 for more details). Specifically, we found that a correlation exists between students' abilities to mediate intuitive thinking by reasoning more analytically, as measured by the CRT, and student performance in physics, as measured by the Force and Motion Conceptual Evaluation (FMCE) [50]. This finding is consistent with the results by Wood *et al.,* who also found a similar relationship between CRT scores and performance on the Force Concept Inventory [37]. Perhaps more importantly, however, our analyses also revealed that the correlation between student performance on the FMCE and cognitive reflection skills becomes significantly stronger as student understanding improves over the course of the semester. The latter result is entirely consistent with Stanovich's construct of the mindware gap [30, 49]. We hypothesize that these results can be extended to specific contexts. In Chapter 3 we investigate the impact of cognitive reflection on student performance in the context of Newton's third law.

Consider the mechanisms that have been presented so far. There are several difficulties which a reasoner may need to overcome to successfully navigate a problem. The first is to avoid biases and overcome any incorrect heuristics. The interplay between process 1 and process 2 show that if process 1 suggests a solution, process 2 can intervene and reject that solution. The bat and ball problem is an example of a situation in which many people encounter incorrect heuristics. With a difficulty of this kind, we need to reflect on our intuition and 'catch' our mistake before giving an answer, keeping in mind that process 2 is also subject to biases. Closely related to this idea is the difficulty of retrieving relevant information. Relevant knowledge must be highly accessible (*i.e.,* easily recalled) for it to have a high probability of being considered by a given student. Furthermore, if an explanation is highly accessible and offered first (by contextual cues or heuristics), a reasoner will be less likely to consider alternatives. The final difficulty also has to do with accessing relevant information; however, the difficulty is not of reasoning, but content knowledge. A reasoner must have sufficient mindware to solve the problem. If the reasoner is unaware, or has not been taught the correct method or rules, they will be unable to reason productively. As such, it is imperative in studies that investigate student reasoning to utilize research based instructional materials to help ensure that a specific reasoning approach employed by a student is not likely due to a lack of relevant formal knowledge and skills. In other words, it is important to disentangle mindware from reasoning.

### 1.6. Dissertation summary

The following chapters demonstrate our use of dual-process theories of reasoning as a lens for interpreting observed patters of student reasoning. Chapter 2 investigates student intuitive reasoning approaches in the context of buoyancy. This study was published in *Phys. Rev. Physics Education Research* [38]. Chapter 3 investigates the relationships between cognitive reflection and student performance on Newton's third law problems. This study was submitted to *Phys. Rev. Physics Education Research* in July 2018 and is currently under revisions. Chapter 4 details a study investigating student ability to override intuition in a context which includes multiple sources of intuitive ideas: namely everyday experiences and mathematics. A more detailed summary of each study is presented below.

#### 1.6.1. Chapter 2

As outlined above, a growing body of scholarly work indicates that student performance on physics problems stems from many factors, including relevant conceptual understanding. However, in contexts in which significant conceptual difficulties have been documented via research, it can be difficult to pinpoint and isolate such factors because students' written and interview responses rarely reveal the full richness of their conscious and, perhaps more importantly, subconscious reasoning paths. In the investigation presented in Chapter 2, informed by dual-process theories of reasoning and decision-making as well as the theoretical construct of accessibility, we conducted a series of experiments in order to gain greater insight into the factors impacting student performance on the "five-block problem." This problem has been used in the literature to probe student thinking about buoyancy. In particular, we examined both the impact of problem design (including salient features and cueing) and the impact of targeted instruction focused on density-based arguments for sinking and floating and on neutral buoyancy. In this study, we found that instructional modifications designed to remove the strong intuitive appeal of the first-available response led to significantly improved performance, without improving student conceptual understanding of the requisite buoyancy concepts. As such, our findings represent an important first step in identifying systematic strategies for using theories from cognitive science to guide the development and refinement of research-based instructional materials. This study was published in *Phys. Rev. Physics Education Research* in March 2018.

### 1.6.2. Chapter 3

The study presented in this chapter demonstrates the relationship between student cognitive reflection skills and performance on Newton's third law problems. This study was motivated by the observation that after targeted instruction designed to improve student conceptual understanding of physics, a significant fraction of students are not able to answer many questions in a consistent manner. Prior research suggests that even those students who demonstrate that they acquired the relevant knowledge and skills (*i.e.,* possess the requisite "mindware") still tend to rely on their intuitively appealing (and often incorrect) ideas. This study aimed to provide insights into cognitive mechanisms that may lead to the identified inconsistencies in student reasoning. We present results of an empirical investigation guided by the Dual Process Theories of Reasoning and accompanying theoretical constructs of cognitive reflection and "mindware." Specifically, a set of hypotheses was proposed to establish a link between student abilities to mediate intuitive responses by engaging the analytical processing in a more productive manner (*i.e.,* cognitive reflection) and performance on physics questions that elicit strong intuitive responses. The Cognitive Reflection Test (CRT) was used to measure students' cognitive reflection skills. A set of screening-target questions in the context of Newton's third law was developed to assess student reasoning approaches in physics. Results suggest that, in the presence of the necessary "mindware," those students who possess a higher level of cognitive reflection skills are more likely to (1) arrive at a correct answer on a question that tends to elicit strong intuitive, but incorrect response; (2) recognize the need for justifying their answers, even if a correct answer does not require rejecting an intuitively appealing response; and (3) engage in consistent reasoning.

### 1.6.3. Chapter 4

In the investigation, we implemented web-based lab activities in order to gain greater insight into the factors impacting student performance in the context of mechanical waves. Instructional materials and practices for mechanical waves have undergone extensive studies and revisions, yet student intuitive responses remain difficult to influence. In this study, we examined student responses on problems that contained the multivariate expression $v = \lambda f$. We attempted to influence intuitive "everyday" approaches and mathematical intuition with the inclusion of highly relevant video simulations of the physical situation at hand. The investigation found that instructional feedback designed to impact incorrect physical concepts without explicitly addressing mathematical intuition produced some conceptual gains, but both intuitive conceptual ideas and particularly intuitive mathematical approaches remained prevalent.

# 2. PROBING STUDENT REASONING APPROACHES THROUGH THE LENS OF DUAL-PROCESS THEORIES: A CASE STUDY IN BUOYANCY

## 2.1. Introduction

In this study, we use the dual-process theories of reasoning and decision-making as a lens for interpreting, in a mechanistic fashion, the observed patters of student reasoning. The topic of buoyancy is used as a context for this investigation. We propose and test several hypotheses that probe the impacts of specific factors and instructional circumstances on student tendencies to engage in the identified reasoning approaches. Finally, we describe instructional modifications, informed by research, that appear to produce positive shifts in productive reasoning in the specific context of buoyancy. The generalizability of our results as well as further directions for research and curriculum development are also discussed.

For the purpose of curriculum development, we are mainly concerned with improving student ability to recognize and apply relevant information. As such, in this study, we use the framework of accessibility. Specifically, we aim to explore the utility of dual-process theories of reasoning (DPToR) and the theoretical idea of accessibility as a guide for interpreting student reasoning in physics and for designing interventions aimed to help students reason more productively. We propose and test several hypotheses motivated by both empirical observations of student reasoning and accessibility, which acts as a mechanism to initiate a particular reasoning path. The accessibility of an idea is influenced by many factors, including the salience of problem features and prior

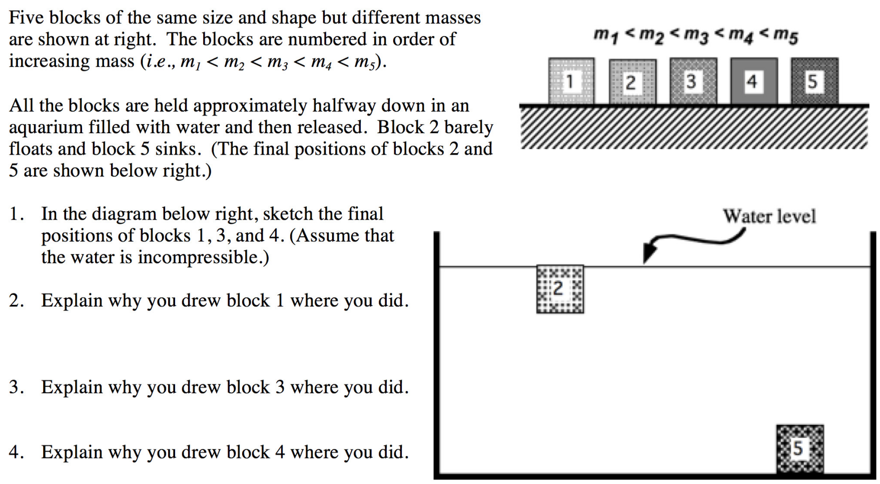



Figure 2.1. The original 5-block problem.

knowledge and experiences. In this study, we operationally define an idea to be more accessible if (1) it is cued through surface features of a task or it appears to be more familiar through prior knowledge or an explicitly stated rule and (2) the application of the idea results in a shorter path to the answer (*i.e.*, fewer steps) or requires rather straight forward "rule-based" reasoning (*e.g.*, if . . . , then. . . ). It is important to note that we do not measure the accessibility of an idea directly. Instead, we use this theoretical framework to propose a hypothesis. We then infer the accessibility of an idea based on the relative prevalence of the idea in student responses. In particular, we argue that many individual students in each class have more or less the same set of available ideas and that, if we assume that each student's response is a reflection of which of those ideas is most accessible for him/her, then the prevalence of particular responses in the group is a measure of which ideas are most accessible.

### 2.1.1. Motivation for 5-block buoyancy problem as a reasoning context

The "5-block" buoyancy problem was designed and extensively used in an investigation of student understanding of buoyancy conducted by Loverude *et al.* The researchers identified a number of persistent difficulties with concepts and principles related to sinking and floating and then used the result of their investigation to develop a tutorial on buoyancy included in *Tutorials in Introductory Physics* by the Physics Education Group at the University of Washington (UW PEG) [6, 7, 51].

On the 5-block problem, students are asked to consider five blocks of the same size and shape, but different mass, as shown in Figure 2.1. In the diagram, the blocks are numbered in order of increasing mass. The students are told that the blocks are held halfway down in the tank and then released. The final positions of blocks 2 and 5 are given. The students are asked to sketch the final positions of the remaining blocks and to explain their reasoning. In order to answer correctly, students may take two approaches.

- The *forces argument* requires students to recognize that, at the position shown, the buoyant force on block 2 is approximately equal to the weight on the block ($m_2g$). Students then must compare the buoyant forces on each block just after it is released (also equal to $m_2g$) to the weight of the block. This approach is illustrated by the following student response: *"[Block 3] is heavier and more dense than 2. If 2 is barely floating and 3 has a greater gravitational force than 2 and the same buoyant force, it will have a net force down."* By similar reasoning, Block 1 floats while block 4 sinks.

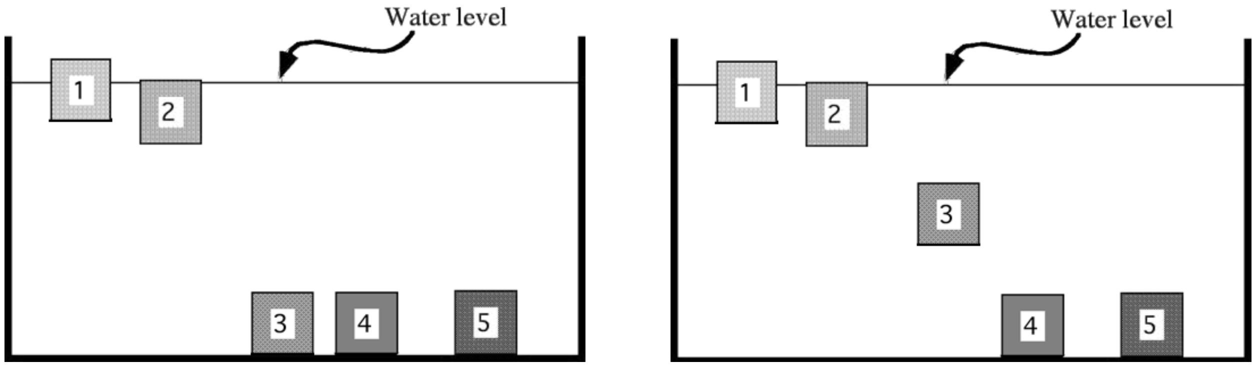



Figure 2.2. Responses to the 5-block problem. Considered to be correct.

- The *density argument*, on the other hand, requires students to recognize that because block 2 "barely floats" the density of the block is approximately equal to that of water. Students must compare the densities of each block to that of block 2 (and therefore to that of the water) in order to predict the sinking or floating behavior. Since the density of block 1 is less than that of block 2, block 1 will be floating as shown in Fig 2.2. By similar reasoning, the densities of blocks 3 and 4 are greater than that of block 2, and blocks 3 and 4 must therefore sink. One example of a student response to the placement of block 3, also considered to be correct, is illustrated by the quote: *"Since the densities are increasing by unknown increments, one can assume 3 is either the exact same density as water, leaving it suspended or more dense, causing it to sink."*

The most common incorrect student response involves the blocks floating in a "descending line": block 1 floats higher than block 2, block 3 somewhat lower than 2, and block 4 somewhat lower than block 3 (See Fig. 2.3). Neither 3 or 4 are on the bottom of the tank; instead, they are suspended near the middle. We hypothesized that responses of this nature may not necessarily be due to shortcomings in student conceptual understanding. We argue that some students may not even attempt to reason with formal knowledge acquired as a result of instruction. Instead,

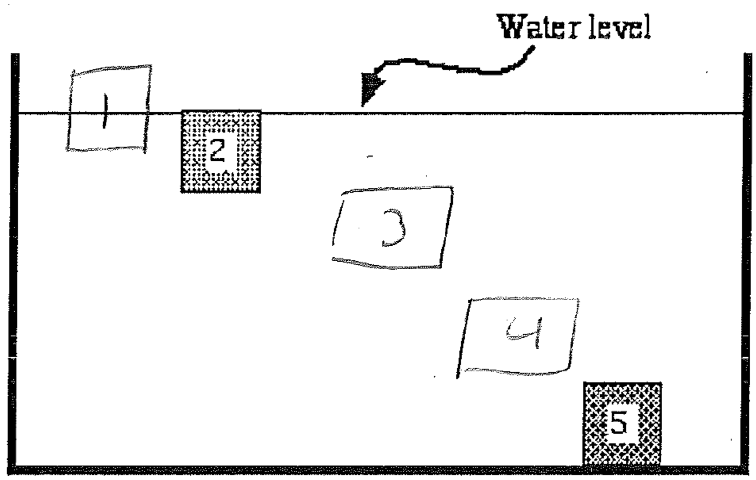



Figure 2.3. Incorrect, *descending line* response.

they may engage in intuition-based arguments cued by specific features of the questions or by their perceptions of how things work in everyday life. The possibility that many incorrect responses to the 5-block problem "were attempts to justify a prediction based on intuition" was also suggested by Loverude *et al.* for further investigation [6, 52]. In the current study, we build on this prior research in order to examine the origins of relevant student intuitions and to probe the impacts of such intuitions on student reasoning.

## 2.2. Instructional context and study design

### 2.2.1. Original buoyancy tutorial

The buoyancy tutorial designed by UW PEG focuses on student understanding of forces and pressure, and on the application of these ideas to investigate the phenomenon of buoyancy [7]. The tutorial was explicitly designed to help students develop a robust understanding of the buoyant force in the context of sinking and floating. The idea that the density of an object could also be used as a predictor of sinking and floating was not explicitly discussed in the tutorial activities, but was included in the homework. A brief description of the tutorial activities is included in Appendix A.

### 2.2.2. Instructional context

This study was conducted in an introductory calculus-based physics course for science and engineering majors at North Dakota State University (NDSU). Over the five semesters of the study, the course was taught by two different instructors, both of whom implemented active-learning techniques in a large-enrollment lecture environment (N$\sim$60 students). The majority of the students were also enrolled in a weekly, fairly traditional, two-hour laboratory; however, student difficulties with the concepts of sinking and floating were not explicitly targeted during lab instruction. In all semesters, instruction on buoyancy was based on the activities from the UW tutorial. However, various modifications to the original buoyancy tutorial were implemented as part of this study, as described later in this dissertation.

### 2.2.3. Overview of the study

Student reasoning approaches were probed in a three-part investigation.

- In Part 1, student written responses to the 5-block problem were collected and several "think aloud" interviews were also conducted. Reasoning approaches were analyzed through the lens of dual-process theories of reasoning (DPToR) and the framework of accessibility.
- In Part 2, two experiments were conducted in order to probe the impact of changes in the task design on student reasoning.
  - In Experiment 2.1, the 5-block problem was redesigned to reduce the salience of specific surface features thought to cue the descending line response. We hypothesized that these changes would reduce the accessibility of the unproductive "descending line" reasoning path.

  - In Experiment 2.2, the 5-block problem was redesigned to shift the salience from mass to density. We hypothesized that shifting the focus of the original 5-block problem from the ranking of masses of the blocks to the ranking of densities of the blocks would increase the accessibility of the density argument and thus increase the fraction of correct responses.

- In Part 3, two further experiments were conducted in order to probe the impact of targeted instruction on student reasoning approaches.

  - In Experiment 3.1, the original UW tutorial was modified to include explicit density instruction. We hypothesized that modifications that include explicit instruction on density would increase the accessibility of the density argument even further, thereby improving student performance.

  - In Experiment 3.2, further instructional modifications were made in order to address incorrect student intuitions about sinking and floating from everyday life. We hypothesized that such instruction would reduce the appeal of the incorrect first-available model, and thus make it less likely that students would bypass the analytical process. At the same time, we speculated that the increased emphasis on density would increase the likelihood that students' resulting formal reasoning would be based on density, and, given that it is easier to reason this way, improve performance.

While various conditions were tested in each experiment, a sequence of three screening questions was developed and employed as an invariant measure of student thinking throughout the study. The screening questions were designed to probe student understanding of the concepts necessary to arrive at a correct answer to the 5-block problem using the force argument. In addition, comparisons of student performance on the screening questions under different experimental conditions throughout the study helped us ensure that the instructional modifications implemented here did not impair the development of student understanding achieved by the original tutorial. A description of the screening questions is included in Appendix A.

### 2.3. Part 1: Probing student reasoning approaches to the 5-block problem

One of the goals of Part 1 was to analyze student reasoning approaches (through both written exam responses and think-aloud interviews) through the lens of DPToR. The second goal was to establish a baseline performance measure on the screening sequence and the 5-block problem at NDSU.

#### 2.3.1. Methodology

The original UW tutorial was implemented by an instructor in the interactive lecture-based format. The instruction was entirely focused on the buoyant force; no instruction on density as a predictor of sinking and floating was formally provided in the course. The screening sequence and the 5-block problem were administered on the final exam upon completion of the original UW tutorial on buoyancy (N=66). In addition, a number of student interviews were conducted prior to formal instruction on buoyancy in the introductory calculus-based physics course. There was no overlap between students participating in the interviews and answering the questions on the final exam because the interviews and final exam took place in separate semesters. The interviews followed a semi-structured think-aloud protocol during which a single student was presented with the 5-block problem and was asked to predict the positions of the blocks while verbalizing his or her thoughts. In this section, we first present student interviews and then discuss student written responses to the screening sequence and the 5-block problem administered on the final exam.

### 2.3.2. Results

The think-aloud interview responses to the 5-block problem provided evidence that student intuition plays a role in producing the descending line response. Below, we present results from two illustrative student interviews.

*Interview 1:*

[1] S: **[Reading question]** "In the diagram below, sketch the final position of blocks 1, 3 and 4. Assume that the water is incompressible."

[2] S: "Ok, I'm basically just going off of mass and what I know from chemistry about density. Block 1 is lighter than block 2 so it's going to float a little bit more. And then block 3 is heavier than block 2, so it's going to be a little bit lower, and then lower for 4."

[3] I: "Ok, so for the first one, why did you draw block 1 where you drew it?"

[4] S: "Because it weighs less than block 2, and block 2 is barely floating, so I drew block 1 as floating 'more' I guess."

[5] I: "Ok, and what about 3 and 4? Why did you draw them where you put them?"

[6] S: "Umm…that's a great question. I'm just guessing 'cause they weigh more than block 2 but less than block 5? So they'd be somewhere in there? Either that or they'd be at the bottom. Because they'd weigh…I don't know I'm not sure."

[7] I: "So, thinking about the physics behind it, what would cause block 4 to be lower or block 3 to be lower, or for them to be at the bottom, or at the top? How would you go about figuring that out?"

[8] S: "Well, is this like a density thing?"

[9] I: “You tell me.”

[10] S: “Kind of?”

[11] I: “How would thinking about densities help?”

[12] S: “Because if it’s about densities, then if it’s less dense than water, it would be on top, if it’s more dense than water then it would be on the bottom.”

[13] I: “OK sorry, what did you say? Repeat it one more time.”

[14] S: “If it’s less dense than water it would be at the top. If it’s more dense than water it would be at the bottom. I think. Maybe.”

[15] I: “So where does that information come from, or how do you know that?”

[16] S: “Kind of from chemistry a little bit?”

[17] I: “So then basing it off what you just said, can you justify putting blocks 3 and 4 and 1 where you did?”

[18] S: “Not really. Guess I’m changing my mind.”

[19] I: “Ok, changing your mind how?”

[20] S: “Now I want to keep 1 at the top and 3 and 4 at the bottom. Maybe? But I also don’t know if 3 and 4 are more dense than water. But I’m thinking you can kind of assume because 3 is heavier than 2 and 2 is barely floating, but I don’t know if you can actually say that.”

The exchange above suggests that the student immediately recognized the relevance of both mass and density to the outcome of the presented situation (line 2). The student’s first response, however, was entirely based on the ranking of the mass or, more specifically, on the notion that the ranking of the mass predicts an identical pattern in the ranking of depth. In fact, the student sketched the descending line of blocks right away, before attempting to articulate her reasoning or any kind of rule. This observed behavior is consistent with DPToR, which suggests that, in many situations, an answer comes to mind first and only then an argument is built to support what is already believed to be true. In addition, the student seemed to view mass and density as competing variables, thus failing to recognize that, since the volume of the blocks is the same, both arguments must lead to the same prediction. Even after the relevant density rule surfaced, the student appeared to be struggling with committing to a specific reasoning approach without deciding first what the problem is "about": "[...] if it's about densities, then [...]" (lines 8-12). The student initial response appears to be consistent with the notion that the problem is "about mass."

Note that one of our explicit goals is to view student responses through the lens of DPToR. However, the interview responses above could also be interpreted through other theoretical frameworks such as "framing" or "resources." Framing is based on a learner's perception of "what a problem is about" and is linked to the learner's prior experiences and expectations [14,15]. Similarly, the theoretical framework of "resources or knowledge in pieces" suggests that student selectiveness in choosing an argument is cued by specific contextual features [16, 17]. We argue, however, that the framework of the DPToR encompasses the theoretical underpinnings of both "framing" and "resources" [1, 11]. Through the lens of the DPToR, the subconscious and automatic selection of a "frame" or a particular "resource" occurs through the heuristic process based on prior experiences, contextual cues, expectations, and other factors.

*Interview 2:*

[1] S: **[Reading question]** "In the diagram below, sketch the final position of blocks 1, 3 and 4. Assume that the water is incompressible."

[2] S: "Alright, so for block one I would think it would be above block 2 because it weighs less. It has less gravitational force on it. So I would put it above block 2 somewhere. Then, we know that block 5 is at the very bottom of the pool, so 3 and 4 are probably in between them [blocks 2 and 5]. 4 weighs more than 3 so I think 4 is probably going to go further down and could be at the bottom. We don't know necessarily, but I would assume it's not."

*Post-interview, after answer has been discussed:*

[3] I: "What did you think about [the 5-block] problem?"

[4] S: "I thought it was... I mean I didn't really know anything about the buoyancy force, so it was a little unfair. I thought it was really interesting though, because you wouldn't think that every block would sink the same even if it had different masses, so you'd think they'd be in different positions in the water, but they aren't. I guess I kind of thought of submarines and how it's harder to go down farther."

The second interview provides further evidence that a direct mapping of the ranking of the masses to the ranking of depths is readily available and appears to be subconsciously appealing. In addition, the interview debriefing revealed that this notion may be rooted in student observations that floating at different levels underwater is a common everyday occurrence, given the behavior of submarines and fish. It appears that without a robust understanding of how real entities (*e.g.,* submarines) maintain neutral buoyancy, students may be surprised that in the case of ordinary objects (*e.g.,* solid blocks), achieving neutral buoyancy is a very rare phenomenon.

Student written responses to the screening sequence and the 5-block question administered on exams provide further evidence that intuition plays a significant role in producing the descending line response. Post-instruction, the majority of students were able to answer each of the questions in the screening sequence with 83%, 66%, and 65% correct, respectively, on the three questions. Only 44% of students correctly answered all three screening questions with correct and complete reasoning, thereby demonstrating that they possessed the formal knowledge and skills necessary to make a correct prediction on the 5-block problem using the force argument. Slightly less than two-thirds (62%) of all students answered the 5-block problem correctly with complete reasoning. Of these students, 62% applied the force argument, while the rest (38%) used the density argument, as shown in Table 2.1. It is important to note that a considerable fraction of the students utilized the density argument in arriving at the correct prediction even though no explicit instruction on density as a predictor of sinking and floating was included in the tutorial. This pattern is consistent with the results reported by Loverude *et al.* [7, 52]. Moreover, a large fraction of students (∼30%) who correctly applied conceptual knowledge of the buoyant force to the screening questions abandoned this line of reasoning in favor of the density-based argument on the 5-block problem. Through the lens of DPToR, student use of the density argument without explicit instruction suggests that this rule-based argument may be more readily available (*i.e.,* accessible) to students than the multi-step force argument. As a result, the reasoning process of some students may start along the dimension of density as opposed to forces.

Further analysis of student reasoning approaches showed that about 20% of those students who used correct reasoning on all of the screening questions provided the descending line response on the 5-block problem. The following student response illustrates this inconsistency in reasoning.

Table 2.1. Student performance on the screening sequence and versions of the 5-block question over the five semesters of this study.

| **Experiment** | **Modification to original UW Tutorial** | **Performance on screening sequence (percentage correct)** | **Version of 5-block question** | **Performance on 5-block question (percentage correct)** | **Percentage of correct responses supported by a density argument** |
|---|---|---|---|---|---|
| Part 1: (N=66) | None | 44% | Original | 62% | 38% |
| Part 2: | | | | | |
| Experiment 2.1 (N=60) | None | 38% | Surface features are modified | 50% | 19% |
| Experiment 2.2 (N=49) | None | 39% | Blocks are ranked by density instead of by mass | 51% | 88% |
| Part 3: | | | | | |
| Experiment 3.1 (N=86) | Density additions | 35% | Original | 57% | 65% |
| Experiment 3.2 (N=72) | Density and neutral buoyancy additions | 35% | Original | 74% | 15% |

On the screening sequence, the student correctly sketched a free-body diagram and provided the following answers to the remaining two questions:

**Screening Question 2 Response:**

*"The buoyant force on block A is less than the magnitude of the buoyant force on block B. Each block has the same volume and mass, but the volume displaced by block two [sic] is greater (Archimedes' principle)."*

**Screening Question 3 Response:**

*"The magnitude of the buoyant force on block B is equal to that of block C because they both displace the same amount of water."*

On the 5-block question, however, the student did not apply the same relevant concepts. Instead, the student sketched the descending line of blocks and stated the following when justifying the locations of the blocks:

*"Block 1 is the lightest. Since Block 2 barely floats, it is safe to assume that Block 1 will be floating more. Block 3 is less than Block 2 but greater than 4 and 5, so putting block 3 submerged in the water, but closer to the surface than the bottom is a good assumption. Block 4 is just a little lighter than block 5, so we can assume that it is very close to the bottom of the water, but not touching the bottom."*

It appears that the student did not attempt to apply the same formal reasoning that was applied to the screening questions. Instead, the student reasoning may have been cued by the provided ranking of the blocks according to mass. Since the intuitive notion that the lightest block 1 will be floating on the surface is consistent with the formal reasoning, the positions of blocks 1, 2, and 5 might strongly cue the idea that the ranking of the depths of the blocks is determined by the ranking of the corresponding masses. As such, according to DPToR, a student may not even engage the analytical process in order to check for consistency between the intuitively strong notion of the descending line of blocks and the formal knowledge of buoyancy.

A similar pattern of responses was exhibited by other students who correctly answered the screening questions, but failed to make correct predictions on the 5-block problem. These responses show that despite student ability to systematically analyze the three screening questions and to articulate correct understanding of the relevant concepts, one in every five students did not apply that same understanding on the 5-block problem; this is consistent with results obtained in other contexts [10,11]. Much like in the student response discussed above, many responses did not provide any evidence that the students had attempted to apply relevant concepts to the 5-block problem. We speculated that the "descending line" response is likely due to features of the questions cueing this intuitive, highly plausible (but incorrect) response. The high accessibility of such a compelling response may result in students feeling that formal support for their answers is unnecessary because the predicted outcome is "intuitively obvious."

### 2.3.3. Summary

After instruction, written exam responses fell into one of the three categories: (1) correct answers supported by the density argument, (2) correct answers supported by the force argument, and (3) incorrect descending line responses, with approximately five percent of responses not fitting into any of these three categories. Data indicate that even students who demonstrated conceptual understanding of the buoyant force on the screening questions did not always employ this understanding on the 5-block problem. Instead, many of the students used the (correct) density argument even without explicit instruction, while another significant fraction of the students seemed to rely on the incorrect but intuitively appealing notion that the ranking according to mass is directly mapped to the ranking according to depth. We hypothesized that such inconsistencies in student reasoning approaches are due to the heuristic process subconsciously choosing a more accessible reasoning path on the 5-block problem, typically involving either density or a mapping of the mass ranking rather than an application of the buoyant force (which seems to be less readily available). Responses involving a mapping of mass to depth are usually not justified by formal knowledge developed as a result of instruction. Instead, such responses may be rooted in incorrect student perceptions about sinking and floating in the real world, and students may therefore fail to perceive a need for any kind of rigorous justification of their thinking.

## 2.4. Part 2: Impact of changes in the problem design on student reasoning

### 2.4.1. Experiment 2.1: Modifications of the 5-block problem by removing prominent cuing features

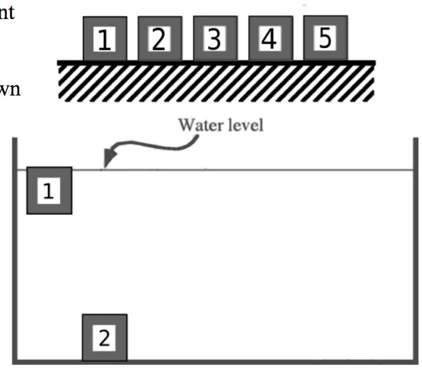
Prior research has suggested that the salient surface features of a task may distract from formal reasoning paths and cue intuitive responses [10, 11, 19, 20]. We hypothesized that removing specific distracting features from the 5-block question may decrease the accessibility of the "descending line" response and result in improved student performance on the modified questions. In the new version of the 5-block problem shown in Fig. 2.4: (1) the ascending ranking of the masses was removed, (2) the blocks and their relative masses were introduced one-by-one to give students an opportunity to consider each block individually, (3) the two blocks with the known sinking/floating behavior were placed next to each other in order to eliminate the space for a "fill in the blanks" descending line of blocks. The task of sketching the final positions of the blocks remained the same and the sequence of screening questions remained unchanged.

### 2.4.2. Results

Student performance on the screening questions showed that approximately equal fractions of students were able to answer the set of screening questions correctly in both administrations, as shown in Table 2.1. Analysis of student responses on the two versions of the 5-block problem also did not reveal any significant differences in student performance: (50% correct in Experiment 2.1 vs. 62% correct on the original problem; Fisher Exact $p = 0.2$). As in Part 1, even those students who demonstrated that they possessed formal knowledge of the buoyant force on the screening questions did not apply this knowledge on the modified 5-block problem (35% gave the descending line response after answering all screening questions correctly). These results suggest that the removal of distracting features of the problem, particularly those hypothesized to elicit and/or facilitate a descending line response, did not appear to mitigate the intuitive appeal of that response for some students.

Five blocks of the same size and shape but different masses are shown at right.

All the blocks are held approximately halfway down in an aquarium filled with water and then released. Block 1 barely floats and block 2 sinks. (The final positions of blocks 1 and 2 are shown below right.)




1. Consider block 3. The mass of block 3 is *greater than* the mass of block 1 and less than the mass of block 2. (*i.e.* $m_1 < \boldsymbol{m_3} < m_2$).

   Sketch the final position of block 3 on the diagram above and explain why you drew block 3 where you did.

2. Consider block 4. The mass of block 4 *less than* the mass of block 1. (*i.e.* $\boldsymbol{m_4} < m_1 < m_2$)

   Sketch the final position of block 4 on the diagram above and explain why you drew block 4 where you did.

3. Consider block 5. The mass of block 5 is *greater than* the mass of *both* block 1 and block 3, but *less than* the mass of block 2 (*i.e.* $m_1 < m_3 < \boldsymbol{m_5} < m_2$).

   Sketch the final position of block 5 on the diagram above and explain why you drew block 5 where you did.

Figure 2.4. The modified version of the 5-block problem in Experiment 2.1: Removing salient distracting features.

### 2.4.3. Experiment 2.2: Modifications of the 5-block problem by shifting focus from mass to density

In Experiment 2.2, students considered a question sequence identical to the original 5-block problem except the ranking of the masses was replaced by a ranking of the *densities*, as shown in Fig. 2.5. This modification was motivated by our prior results suggesting that the accessibility of the density argument for this student population already appeared to be high. We argued that the density argument could be categorized as "rule-based": if the density of an object is higher than that of water, the block sinks; if it is lower, the block floats. This rather straightforward reasoning path (*i.e.*, if. . . , then. . . ) may be more accessible for some students than the force argument, which does not follow an algorithmic path. As such, we hypothesized that by cueing the density of the blocks explicitly, we may increase the accessibility of the density argument even further, which may therefore increase the likelihood of students making correct predictions regarding the sinking and floating behavior of the blocks.

### 2.4.4. Results

Data analysis revealed no significant impact of the task modification on student performance: 39% of the students were correct on all screening questions, with 51% correct on the modified 5-block problem. However, an important (but perhaps not surprising) difference in the response pattern was observed: 88% of students who answered the modified 5-block problem correctly used the density argument, while the remaining 12% used the force argument. The latter responses explicitly referred to the greater mass and equal buoyant forces. These results suggest that the modified 5-block problem may have increased the accessibility of the density argument for some students. At the same time, it did not remove the intuitive appeal of the descending line of blocks argument for a large fraction of students.

### 2.4.5. Summary

The prevalence of the descending line responses does not appear to be impacted by altering the prominent features of the 5-block problem. In addition, even without explicit instruction, the correct density argument is common and it becomes even more prevalent when density is explicitly included in the prompt. Interpretation of these results through DPToR suggests that by altering prominent features of the task, we were able to increase the accessibility of a specific correct reasoning approach (*i.e.,* the prevalence of the density argument), but failed to decrease perceived relevance of the incorrect descending line approach enough to see a noticeable increase in correct responses. As such, our initial hypothesis that the heuristic process cues the persistent descending line response based on the features of the task does not find substantial support in the collected data. Specifically, it seems that while student *arguments* can be affected by changing the prompt, increasing the *performance* on the problem may require modification to instruction. As a result, we then decided to explore an additional set of hypotheses related to relevant instruction.

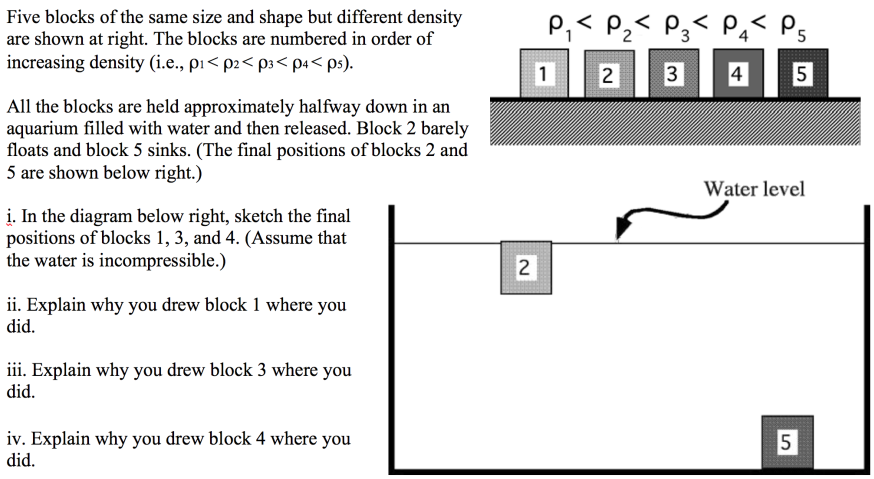



Figure 2.5. The modified version of the 5-block problem in Experiment 2.2: Ranking of the mass is replaced by ranking of density.

### 2.5. Part 3: Impact of instruction on student reasoning and performance

We tested the impact of two sets of instructional modifications designed based on the following justifications. In Experiment 3.1, we drew on research by Heckler and colleagues that examined the impact of competing variables on student reasoning. They discovered that (1) providing students with a correct rule for predicting an outcome reduces reliance on an irrelevant competing variable, and (2) providing opportunities for practice decreases the response time, thereby promoting the likelihood that a specific desired argument comes to mind faster. These experiments were performed in a laboratory setting and the interventions were given just before students were asked to consider a task. In our case, we probed the impact of instructional modifications on the accessibility of the density argument. We speculated that the rule-based density argument may have more perceived usability than the forces argument due to its (relatively) straightforward application. This in turn may make the density argument more competitive with the intuitive "descending line" idea. As such, we included supplemental exercises that (1) introduced the concept of density as a predictor of sinking and floating and (2) strengthened the links between the force and density arguments. We hypothesized that explicit instruction on density will further increase the accessibility of the density argument and will therefore improve student performance on the original 5-block problem.

In Experiment 3.2, we tested an alternative hypothesis. The results from Parts 1 and 2 suggest that changes in the task design were impacting the accessibility of the correct density argument while failing to reduce relevance of the incorrect descending line approach. This suggests that the students' "descending line" responses may be based on strong, everyday perceptions of sinking and floating behavior, as illustrated by student interview 2 in Part 1. These students may incorrectly think that regular objects free-floating under water are a common occurrence. This perception may be rooted in student observations of fish, submarines, and other objects and animals that can adjust their average densities so as to become neutrally buoyant. For these students, the outcome of regular objects (such as solid blocks) floating at different depths below the water surface may be perceived to be common and therefore highly plausible. While formal instruction on buoyant forces may have helped these students develop the conceptual understanding necessary to answer the 5-block problem correctly, these students may still have perceived little need to apply such understanding to justify predictions based on their intuitive ideas. If, indeed, the descending line responses are strongly cued by student intuitions about how "real world" works, then the instructional modifications proposed in Experiment 3.1 are also not likely to produce significant shifts in student performance. We hypothesized that for the students who answer the 5-block problem correctly, the plausibility of objects free-floating under water may not be as strong. Therefore, these students may be more likely to apply formal reasoning to arrive at a correct answer. As such, for this category of student thinking, the inclusion of explicit instruction on density is likely to further increase the accessibility of the density argument and, as such, increase the prevalence of this argument on the 5-block problem. (This is consistent with the shifts in the prevalence of the density arguments in Experiments 2.1 and 2.2.)

Finally, we hypothesized that if the nature of the descending line responses is linked to student perceptions about the behavior of fish and submarines, then, in order to decrease the accessibility of such responses, further instructional modifications designed to highlight the complex sequence of steps required to achieve neutral buoyancy at different depths are required. In order to test this hypothesis, in Experiment 3.2, we modified instruction by implementing a set of questions designed to guide students through the reasoning steps necessary to understand the conditions that allow a submarine to float under water and to change its depth. We also included hands-on, in-class activities that allowed students to experience firsthand the sinking and floating behavior of common objects. We hoped that the students would realize that it is extremely unlikely to observe common objects floating under water and it is also extremely difficult to achieve neutral buoyancy in common laboratory settings. We hoped that these experiences would help students recognize that the descending line outcome of the 5-block problem is highly unlikely, which in turn may prompt students to consider alternative outcomes by applying formal reasoning approaches.

### 2.5.1. Experiment 3.1: Explicit instruction on density as a predictor of sinking and floating

#### 2.5.1.1. Supplemental exercises focusing on density as a predictor of sinking and floating

Two exercises were designed to supplement the development of ideas in the original UW tutorial. The first exercise was introduced immediately after the discussion of the buoyant force. Students were asked to consider three objects of the same shape: two blocks of different density and a cubical volume of water with an imaginary boundary around it. Students related the sinking and floating behavior of the blocks to the density of the blocks relative to that of the water. The students were then asked to articulate a simple “rule” for predicting whether an object will sink or float based on its density. A second supplemental exercise was included at the end of the tutorial. The students were asked to predict the sinking or floating behavior of a block and support their predictions by applying both the force and the density arguments.

### 2.5.2. Results

The screening sequence and the original 5-block problem were used to examine the impact of the instructional modifications designed to increase student accessibility with relevant concepts of density and buoyant force. Analysis of student responses did not reveal any significant improvement in student performance. However, as in Experiment 2.2, the prevalence of the density-based correct responses increased substantially compared to that in Part 1. Respectively, 62% and 57% of students answered the original 5-block problem correctly in Part 1 and in Experiment 3.1. Of these correct responses, 38% used the density argument in Part 1 compared to 65% in Experiment 3.1. (See Table 2.1). This suggests that, as a result of instruction, the accessibility of the density argument was increased, for at least some fraction of students. At the same time, for those students who answered the 5-block problem incorrectly, the possible increase in the accessibility of the density argument was not enough to override the intuitive appeal of the descending line prediction. This result is consistent with the alternative hypothesis above used to justify Experiment 3.2.

### 2.5.3. Experiment 3.2: Including hands-on and real world examples in instruction

In this study so far, we have shown that changing the design of the 5-block problem did not significantly impact student performance, thereby suggesting student intuition is not cued by salient features in this particular case. We have also shown that instructional changes targeting the density argument did not significantly improve performance. Evidence from student interviews suggests that students tend to draw upon intuition involving real world situations (submarines, fish, *etc.*) in which "floating in the middle" is common. This is further supported by Heron *et* *al.* who reported that significant improvement on the 5-block problem was achieved only after implementation of lengthy experimental lab activities that included instruction on density [7, 52]. In Experiment 3.2, we modified the instructional sequence used in Experiment 3.1 in order to: (1) include explicit hands-on activities to help students recognize that neutral buoyancy is difficult to achieve, (2) give students opportunities to analyze conditions necessary for objects of variable average density (*e.g.,* submarines) to achieve neutral buoyancy, and (3) help students articulate the steps necessary for a submarine to move between different depths. We hypothesized that these activities, combined with the density instruction implemented in Experiment 3.1, would decrease the accessibility of the descending line "floating in the middle" argument, which, in turn, would yield improved student performance on the 5-block problem.

#### 2.5.3.1. Supplemental exercises focusing on hands-on activities and real-world applications

The hands-on activities in the modified tutorial followed a sequence of intellectual steps nearly identical to that contained in the sections of the original tutorial introducing the buoyant force and focusing on sinking and floating of hypothetical solid blocks. However, the modified tutorial included simple experiments instead of considering hypothetical sold blocks; Students were provided with two small test tubes, identical in shape, and a large beaker of water. The tubes were partially pre-filled with different masses and a small amount of water. Students observed the motion of the tubes released under water and analyzed the forces acting on the tubes. Tube 1 floated easily above the surface, while tube 2 was very nearly neutrally buoyant and rose to the surface slowly. Students observed the sinking or floating behaviors of the tubes, noted the differences in their motions, and analyzed forces acting on the tubes right after they were released under water. Then, the students were asked to predict and analyze the motion of tube 2 after more mass was added to the tube. Students were given opportunities to check their predictions. This sequence of hands-on activities bore no resemblance to the 5-block problem.

After students analyzed the motions of the tubes using both the force and the density arguments (similar to the tutorial version in Experiment 3.1), they focused on motion of objects of variable average density, such as submarines. Students were guided to articulate conditions necessary for a submarine to remain under water at rest and to move between different depths. Students were asked to use both the force and the density arguments. Finally, a hypothetical dialogue between two fictitious students was presented, which concerned two submarines that were identical in shape but floating at two different depths. Student 1 incorrectly states that the mass of the submarine at a shallower depth is less than that of the submarine deeper under water. Student 2 agrees and supports this incorrect prediction by stating that the ballast will be lower for the submarine closer to the surface. Students were expected to recognize the incorrect elements of reasoning presented in the dialogue and to justify their answers. Specifically, students must recognize that, since the buoyant force does not depend on depth, the ballast at either depth must be identical to ensure that the net force on each submarine is zero. As such, both students in the dialogue incorrectly related the mass of the submarine to the depth.

### 2.5.4. Results

The screening sequence and the original 5-block problem were again used in order to examine the impact of the instructional modifications. Performance on the screening questions was essentially unchanged, but the fraction of correct responses on the 5-block problem increased significantly when compared to experiments 2.1, 2.2, and 3.1 (74% correct with correct reasoning; Fisher Exact $p < 0.05$). While the performance was not significantly different compared to that in Part 1, the increase in performance in Experiment 3.2 is statistically significant compared to all other semesters. This suggests that perhaps student performance in Part 1 is an outlier. Moreover, for the purposes of this study, the increase in student performance from Experiment 3.1 to Experiment 3.2 is of particular importance as the instructional materials for Experiment 3.2 were modified from those used in Experiment 3.1 as described above. This seems to suggest that the additional instruction on real-world applications with a hands-on component had a significant impact on student reasoning compared to the previous semester. Indeed, some of the student responses to the 5-block problem directly referenced the observations made in class. For example, one student wrote:

> *"It is possible for an object to remain in the center of the aquarium if the forces are just right, however due to the fact that [block] 2 was barely floating and completely submerged and* ***from what I saw with the experiments in class****, I felt that block 3 would sink to the bottom because the buoyant force for blocks 3 and 2 would be equal (same size and both fully submerged). The mass of block 3 is reasonably heavier so the force down by the earth would be great enough for it to sink."*

### 2.5.5. Discussion

Several explanations are possible for the increase in performance on the 5-block problem. We argue that perhaps two factors are critical. First, the hands-on activities provided students with opportunities to observe that it is uncommon for ordinary objects to be in the state of neutral buoyancy. Second, the analysis of conditions necessary for submarines to achieve the state of neutral buoyancy and to move between different depths may have helped address incorrect student perceptions regarding sinking and floating of objects of variable densities. As such, these instructional modifications may have removed the strong intuitive appeal of the descending line response to the 5-block problem. This explanation is consistent with student responses similar to the one quoted above. Moreover, according to DPToR, once the intuitively appealing argument is removed, the remaining possibility becomes most plausible. Since the most plausible answer is now typically consistent with the correct application of the formal knowledge acquired as a result of instruction, students are more likely to provide correct answers with correct reasoning, hence the increase in performance. Alternatively, it could be argued that perhaps the modified instruction is more effective at improving student conceptual understanding of buoyancy or provides more *time-on-task*. However, since the increase in the fraction of correct responses to the 5-block problem occurred without any improvement in student performance on the screening questions (see Table 2.1) and the additional instructional time needed to cover neutral buoyancy was relatively minor, we are inclined to think that the modified instruction achieved exactly what it was intended to do: remove the intuitive appeal of the incorrect descending line response.

The outlined reasoning path suggested by DPToR is consistent with an "answer first, reasoning second" pattern (or confirmation bias). As such, it could be debated whether or not the increase in student performance in Experiment 3.2 is an indication of success. One could argue that although the percentage of correct responses increased, the new reasoning path is no more sophisticated than the one that led to the prediction of the descending line: both start with the most intuitively appealing outcome and lead to an answer. The key, however, is that we were able to isolate, through hypothesis testing, a specific aspect of student thinking that appears to be the source of a very persistent incorrect response shown to be particularly resistant to many research-based instructional efforts to improve student conceptual understanding.

Moreover, according to DPToR, reasoners see the world around them through the lens of the quick and automatic Process 1. It immediately suggests the most plausible mental model of an unfamiliar situation and only *then* the slow and deliberate Process 2 may intervene in order to provide an evaluation. In other words, intuitive thoughts cannot be turned off. Simon defines intuition as "pattern recognition" [53]. While experts possess a large repertoire of experiences that allow them to quickly and effectively recognize both correct and incorrect patterns in their thinking, novice learners lack such experiences. As a result, they are less likely to recognize instances of intuitive thoughts or biased reasoning, even when Process 2 is engaged. We argue that classroom activities that give students opportunities to slow down, examine "first-available responses," and recognize instances of biased reasoning must be viewed as an essential component of instruction that ultimately leads to the development of expertise in physics. This research project advocates for the development of instructional materials that help students refine their reasoning approaches alongside the development of robust conceptual understanding.

### 2.6. Conclusions

This research was motivated by observed patterns of persistent student difficulties that have been shown to be resistant to changes in instruction. We applied dual-process theories of reasoning and decision making and the theoretical framework of accessibility in order to identify factors and instructional circumstances that enhance or suppress productive and unproductive reasoning approaches. The context of buoyancy was chosen as a case study. We hypothesized that the observed patterns of student incorrect responses were more likely to be due to incorrect but intuitively appealing student ideas rather than to a lack of conceptual understanding. Indeed, some of the students who possess the requisite knowledge of relevant ideas may not even attempt to apply this knowledge in order to arrive at an answer. Instead, these students tend to abandon formal approaches to solving a problem if that problem cues a strong intuitive response. Several specific hypotheses were tested in order to pinpoint more precisely the root of student intuitive ideas.

The first set of hypotheses probed whether or not student responses to the 5-block problem are strongly cued by prominent surface features of the task; if so, then deliberate changes in the task design were predicted to alter patterns in student reasoning.

- The Experiment 2.1 hypothesis predicted that prominent surface features of the original 5-block problem strongly cued an incorrect descending line response. To test the hypothesis, several changes were made to the design of the problem; however, no significant differences in student performance were observed.

- The Experiment 2.2 hypothesis predicted that shifting the focus of the original 5-block problem from the ranking of masses of the blocks to the ranking of densities of the blocks will increase the accessibility of the rule-based density argument and thus increase the fraction of correct responses. After the proposed modifications were implemented, no significant difference in the fraction of correct responses was observed. However, the accessibility of the density argument seemed to be increased among those students who provided correct answers to the problem.

The results of testing the first two hypotheses suggested that the observed pattern of incorrect student responses could not be attributed to features of the task design. As such, an additional set of hypotheses was proposed that focused on probing the impact of instructional interventions, rather than the task design, on patterns of student reasoning.

- The Experiment 3.1 hypothesis focused on probing whether or not instructional modifications that include explicit instruction on density will increase the accessibility of the rule-based density argument even further, thereby improving student performance.

- The Experiment 3.2 hypothesis represented an alternative to the Experiment 3.1 hypothesis. It was argued that perhaps the lack of impact of the increased accessibility of the density argument on the fraction of incorrect responses was due to the students' strong intuitive ideas about the presented situation itself. These intuitive ideas about how real world works (*e.g.,* fish swimming under water and submarines moving between different depths) may be producing strong impacts on student reasoning approaches. Specifically, students may be so confident in the descending line response based on their intuitions that they may perceive little or no need to apply formal knowledge acquired as a result of instruction. As such, it was argued that instructional modifications in Experiment 3.1 would produce little or no impact on student performance. The Experiment 3.2 hypothesis predicted that instruction designed to specifically address student perceptions about sinking and floating of objects of variable densities such as submarines would be more likely to produce a positive shift in student performance. It is important to stress that the instructional modifications implemented as part of testing the hypotheses associated with Experiments 3.1 and 3.2 were designed to *supplement* tutorial activities intended to help students develop conceptual understanding of buoyancy. Such understanding is *critical* for helping students recognize shortcomings in their intuitive ideas and for helping students resolve inconsistencies between their predictions and the observed outcomes. While instructional modifications in Experiment 3.1 did not produce a measurable shift in student performance, those in Experiment 3.2 did in fact produce a significant increase in the fraction of correct responses.

We argue that this study illustrates the importance of research aimed at identifying factors and instructional circumstances that enhance or suppress productive and unproductive student reasoning approaches. Dual-process theories of reasoning suggest that the quick and automatic

Process 1 cannot be turned off. As such, instruction that explicitly helps students examine their intuitive ideas is necessary to further improve student performance in physics [54, 55]. However, we argue that the instructional modifications designed and implemented in this study produced an impact solely because they supplemented tutorial instruction already shown to be highly effective at helping students develop a formal understanding of the concepts of buoyancy.

In future work, we plan to examine this argument in more detail. In particular, we suspect that the inclusion of supplemental materials that promote student reflection on intuitive ideas are likely to be more effective if offered as a follow-up to formal instruction aimed at improving student conceptual understanding of relevant topics. In the absence of solid understanding of relevant concepts, a “mindware gap” is produced, which is challenging for the students to bridge in order to productively reconcile inconsistencies in their reasoning [30]. Systematic hypothesis testing could help us ascertain how best to help students navigate their Process 1 responses.

# 3. ESTABLISHING A RELATIONSHIP BETWEEN STUDENT COGNITIVE REFLECTION SKILLS AND PERFORMANCE ON PHYSICS QUESTIONS THAT ELICIT STRONG INTUITIVE RESPONSES

## 3.1. Introduction

Novice physics learners often reveal naive ideas and notions about how "things work" in physics. Some of these ideas are intuitively appealing and persist even after targeted instruction. For example, student predictions that "during a collision, a heavier object exerts a greater force" are very common in an introductory mechanics course, while the notion that "the current is used up" is familiar to experienced instructors teaching electric circuits. Many student ideas of this nature are well documented in the Physics Education Research (PER) literature and appear in nearly every topic across introductory physics (and beyond) [3–5]. However, recent studies revealed that multiple refinements of instructional materials designed to improve student understanding of relevant concepts do not necessarily lead to significant improvements in student performance on tasks that elicit strong intuitive responses. These studies also indicate that a fraction of the students, who demonstrate that they acquired correct formal knowledge and skills as a result of instruction, still tend to rely on incorrect (but intuitively appealing) thoughts [10–13]. In some cases, these students do not find the need to apply formal knowledge to justify their predictions. In other cases, they tend to employ incorrect reasoning to support an answer that they already believe to be correct. The latter is consistent with the "answer first, reasoning next" thinking pattern [10].

The study presented here aims to provide insights into cognitive mechanisms that may lead to inconsistencies in student reasoning. We present results of an empirical investigation guided by the Dual Process Theories of Reasoning (DPToR) and accompanying theoretical constructs of cognitive reflection and "mindware." Specifically, the DPToR, developed and widely used in psychology research, suggest that human cognition relies on two distinct processes: Process 1 (often referred to as the *heuristic process*) and Process 2 (the *analytic process*) [1,2]. The heuristic process is fast, automatic, and subconscious. It is responsible for generating a first available mental model of a situation at hand based on the reasoner's prior knowledge, contextual cues, and overall plausibility. The analytic process, on the other hand, is slow, deliberate, and effortful. In everyday life, first available responses generated by the heuristic process are often described as "first impressions" or "gut feelings." In this study, we refer to such responses as a reasoner's intuition.

Three aspects of the DPToR are particularly relevant to the current investigation. First, reasoners see the world around them through the lens of the fast and intuitive heuristic process. Only after a first available response is generated might the analytic process intervene and provide a thorough evaluation of the validity of that response. As such, according to the DPToR, a reasoner's intuition cannot be turned off. Second, if a reasoner does not perceive any threat to the validity of the first available response, the analytic process is circumvented entirely. This direct reasoning path from the first available response to the final inference is often described as "cognitive miserliness" [48,56,57]. Finally, even when the analytic process is engaged, it operates based on biases of its own. For example, reasoners often exhibit confirmation bias and will build an argument that supports an intuitively appealing response instead of checking for consistency between this response and relevant formal knowledge, or looking for alternative solutions [21, 22].

The mechanistic reasoning paths outlined by the DPToR (and further discussed in Section 3.2) suggest that reasoners who are able to identify their own intuitive thoughts and engage in an effortful, logic-based analytical evaluation of such thoughts are more likely to reason productively. The ability to mediate intuitive thinking by reasoning more analytically is often called *cognitive reflection*. The Cognitive Reflection Test (CRT), developed by Frederick, has been used to gauge this ability [28]. A key feature of the test is that, for many reasoners, each question immediately elicits an intuitively appealing, but incorrect, response. In order to answer correctly, a reasoner must be able to recognize the response as intuition-driven, check the response for validity, and replace it with a correct solution, which is generally obtained without gaining any further knowledge. The following CRT question could be used to illustrate this reasoning path: "A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?" For most people, an intuitive answer of "10 cents" springs to mind; in fact, many quickly embrace this answer as correct without further consideration [1, 28]. Upon reflection, however, one may realize (without necessarily gaining additional insight into the relevant arithmetic) that the correct answer is "5 cents."

The example above also highlights that, for some types of questions, an incorrect response does not necessarily reflect the lack of relevant knowledge and skills, but is rather indicative of reasoning tendencies. Indeed, consistent with the DPToR is an assertion that learners may reason incorrectly because (1) they lack sufficient knowledge of the rules and concepts needed to reason through a given problem and/or (2) they behave as "cognitive misers." In an analogy with computer software, Stanovich refers to the "knowledge of rules and concepts" as "mindware" [30].

While possessing mindware is necessary, it may not be sufficient to answer questions of certain types correctly. Since the first available intuitive response (generated by the heuristic process) appears to be an entry point into a reasoning path, we hypothesized that students who possess a higher level of cognitive reflection skills are likely to be more successful at answering certain types of physics questions correctly and consistently. We argue that taking into account the interplay between intuitive and more deliberate processing, combined with cognitive reflection, may be critical for understanding and addressing persistence of incorrect responses as well as inconsistencies in student reasoning on many physics tasks.

In order to ensure that a specific reasoning approach employed by a student is not due to a lack of relevant formal knowledge and skills, but is rather indicative of reasoning tendencies, we used a "screening-target questions" methodology. Newton's third law was chosen as the context for this investigation due to its ubiquity in physics instruction and a large body of PER literature documenting student thinking in this context. The screening question was designed to probe whether or not a student possesses the relevant "mindware" and is able to apply Newton's third law correctly in a challenging situation that does not necessarily elicit a strong intuitive response. Tasks known to elicit intuitive responses in this context were used to inform the design of "target" questions. Two types of target questions were developed, both involving collisions between objects of unequal masses. Target question 1 elicits an intuitively appealing, but incorrect, response. Target question 2, on the other hand, elicits an intuitively appealing response that happens to be consistent with the formal analysis of the presented situation. We then analyzed, in detail, responses to the target questions from those students who answered the screening question correctly, thereby demonstrating the presence of the relevant knowledge and skills (*i.e.,* the requisite mindware). Three hypotheses were proposed and tested.

***Hypothesis 1.*** Given that students answer a screening question correctly, students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to provide correct answers to target question 1, which tends to elicit an intuitively appealing (but incorrect) response that must be rejected.

***Hypothesis 2.*** Given that students answer a screening question correctly, students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to recognize the need for justifying their answers to target question 2, even if the correct answer does not require a rejection of the first available intuitive response.

***Hypothesis 3.*** Students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to answer the screening-target questions consistently (i.e., provide correct answers to all questions and include correct reasoning on target questions).

Section 3.2 of this chapter provides further discussion of the relevant theoretical background, including interactions between the heuristic and analytic processes, cognitive reflection, and mindware. Section 3.3 includes description of our research methodologies, data collection, and analysis. Our findings and their interpretation are included in Sections 3.4 and 4.5.

### 3.2. Prior research

In the present study, we build on prior research to probe the role of cognitive reflection in student reasoning in physics even further. Specifically, we aim to investigate whether or not, in the presence of relevant mindware, differences in student reasoning approaches could be accounted for by the tendency toward cognitive reflection. We are particularly interested in student reasoning on tasks that elicit strong intuitive responses. Student difficulties with the application of Newton's third law in the context of collisions between two objects of different masses (and/or sizes) are well documented. As such, student reasoning with Newton's third law was chosen as a context for the current investigation.

Work dating back to the early 1980s by Halloun and Hestenes on "common sense concepts" and by Brown on "student preconceptions" documented many student ideas regarding collisions and Newton's third law [58–60]. Hestenes described Newton's third law as "inconsistent with common sense intuitions". He observed that most students adhered to some sort of *dominance principle* according to which a greater mass exerts the greater force or the object causing the motion exerts the greater force on the other object. Brown observed that students have a general view of objects "having" more or less force and being more or less "force-full". Students exhibited particular difficulties with collisions between a moving, massive object (*e.g.,* bowling ball) and a stationary, less massive object (*e.g.,* bowling pin) with only ~5% of students successfully answering after fairly traditional instruction. Maloney investigated "rule-governed approaches" to collisions (and other scenarios) involving applications of Newton's third law [61]. He reported that novice students most commonly argued that (1) mass is the determiner, meaning greater mass produces greater force; (2a) for systems at rest, all forces are equal, but for moving systems the greater mass means greater force; and (2b) for systems at rest, all forces are equal, but for moving systems the 'cause' exerts a greater force. Experienced students, those who previously completed a course in college physics, tended to argue that (3a) for systems at rest or moving at a constant velocity, the forces are equal; for accelerating systems, greater mass exerts greater force; and (3b) for systems at rest or moving at a constant velocity, the forces are equal; for accelerating systems, the 'cause' exerts a greater force.

Not surprisingly, questions involving collisions are included in both concept inventories, the FMCE and the FCI, most commonly used to assess the effectiveness of instruction in introductory mechanics courses [62, 63]. These scenarios have also been utilized to probe various aspects of teaching and student learning. Hammer and Elby used epistemological resources as a framework for gaining insights into student thinking, which subsequently led to the development of instructional materials designed to help students refine their everyday intuition in the context of collisions and Newton's third law [64, 65]. A study by Maries and Singh focused on probing the pedagogical content knowledge of teaching assistants in an introductory mechanics course. The researchers found that some teaching assistants tend to perceive the more massive or more 'active' object as exerting a greater force [66]. They noted that "The fact that even some experts hold this alternate conception after many years of practicing physics points out how strong this alternate conception is and how difficult it is to overcome it in this particular context."

### 3.3. Research methodology

In this section we discuss in detail the "screening-target questions" methodology and describe how prior research informed the development of two specific versions of the collision task that served as target questions in this investigation. Additionally, we discuss the instructional context for this investigation and describe data collection approaches.

#### 3.3.1. Instructional context

This study was conducted in two semesters of a large-enrollment introductory calculus-based mechanics course taught by different instructors. In both semesters, the screening-target questions, discussed in detail below, were administered on a single exam upon completion of all relevant instruction on forces and Newton's laws. Both instructors used research-based, active learning strategies. Materials from the *Tutorials in Introductory Physics* developed by the Physics Education Group at the University of Washington (PEG UW) were implemented in an interactive lecture format [51,67]. Scenarios involving two colliding objects of different masses are not included explicitly into the sequence of UW tutorials on forces. However, both instructors incorporated a multiple-choice "clicker" question featuring a moving heavy truck colliding with a small car. From the choices given, students were asked to determine conditions under which the magnitude of the force that the truck exerts on the car is *greater* than the magnitude of the force that the car exerts on the truck. Students were expected to recognize that, according to Newton's third law, the two forces must be equal in magnitudes under all possible circumstances.

### 3.3.2. Overview of the screening-target questions

This investigation aimed to probe whether or not, in the presence of relevant mindware, differences in student reasoning approaches could be accounted for by the tendency toward cognitive reflection. In order to disentangle (to the extent possible) student knowledge and accompanying skills from reasoning tendencies, we applied the screening-target questions methodology.

***Screening question.*** The screening question, shown in Fig. 3.1, requires students to recognize the applicability of Newton's third law in a situation that is challenging but not associated with a strong intuitive response. Students were asked to consider two identical blocks on a plate lifted upward with the acceleration of $g/2$. Students compared the magnitude of the normal force exerted on block A by block B ($N_{AB}$) to the magnitude of the normal force exerted on block B by block A ($N_{BA}$). In order to answer correctly, students must recognize that the two normal forces constitute a Newton's third law force pair and, therefore, their magnitudes must be equal.

Consider the system shown at right. Blocks A and B each have mass $m$. Plate C, all cables and supports are massless.

The system is lifted with acceleration $g/2$ upward. Ignore air resistance.

Let $N_{AB}$ be the magnitude of the normal force on A by B. Let $N_{BA}$ be the magnitude of the normal force on B by A. Which of the following best represents the comparison of the magnitudes of these two forces?

A. $N_{AB} > N_{BA}$

B. $N_{AB} < N_{BA}$

C. $N_{AB} = N_{BA}$

D. Not enough information to determine.

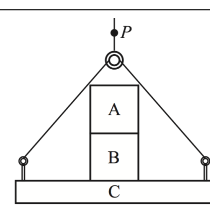



Figure 3.1. 'Screening' question designed to probe whether or not a student is able to apply Newton's third law in a situation that does not elicit a strong intuitive response.

Prior research by Stetzer and PEG UW suggests that, even though this context does not elicit strong intuitive responses of the same nature as those observed with two colliding objects, this screening question could still be challenging for many students. In particular, after traditional lecture instruction, most students do not seem to have significant difficulties with the application of Newton's third law in cases when the system is either at rest or moving vertically with a constant speed (~95% and ~70% correct, respectively). However, when presented with the system accelerated upward, only about 25% of students successfully recognize the applicability of Newton's third law. Perhaps distracted by irrelevant information (*e.g.,* "acceleration is upward"), many students become entangled in more complex lines of reasoning, which typically involve a net force in the direction of the given acceleration, a free-body diagram, and Newton's second law. As such, we argue that a correct student response to the screening question may serve as an indication that the student (1) possesses knowledge of Newton's third law and (2) is able to recognize its applicability in a challenging situation.

***Target questions.*** Two target questions were designed to probe the impact of cognitive reflection skills on student reasoning approaches. Similar to the screening question, the target questions require the application of Newton's third law. However, both target questions feature two canonical characteristics of a collision context known to elicit intuitive responses for some students: (1) one object is more massive than the other, and (2) one object is "active" while the other is initially at rest.

***Target question 1***, *the mine cart* question, is shown in Fig. 3.2. The question involves a moving heavy cart colliding with a light cart initially at rest. Students are asked to compare the magnitudes of the forces that the two carts exert on each other during the collision. For some students, this question elicits a well documented intuitive idea that "during a collision, the force exerted by a heavy active object is greater than that by a light inactive object." As such, in order to answer correctly, these students must, first, reject or suppress this intuitive response, and then replace it with a formal argument based on Newton's third law. This question allowed us to test hypothesis 1: given that students answer a screening question correctly, those students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to provide correct answers to target question 1, which tends to elicit an intuitively appealing (but incorrect) response that must be rejected.

***Target question 2***, *the spaceship* question, is shown in Fig. 3.3. This question also involves a collision between a massive object (a spaceship) and a small moving object (an asteroid). The spaceship is initially at rest, while the asteroid moves with a high speed. The question asks students to compare the *accelerations* of the two objects during the collision. Formal reasoning requires two steps: (1) according to Newton's third law, the forces that the two objects exert on each other during the collision are equal in magnitude; and (2) according to Newton's second law, the acceleration of a light asteroid must be greater than that of the heavy spaceship. As in target question 1, the application of Newton's third law is required to answer this question correctly. (While target question 2 requires one extra step, we did not anticipate significant difficulties with a straight forward application of Newton's second law in this context.) In contrast to target question 1, the correct formal response in this scenario is also consistent with an intuitively appealing answer.

Two identical carts are on a straight, frictionless track. One of the carts (Cart A) is empty; the other (Cart B) is full of heavy rocks. Cart A is initially at rest, and Cart B moves along the track toward Cart A with constant velocity $v_o$. Eventually, Cart B collides with Cart A.


During the collision, is the magnitude of the force exerted on Cart A by Cart B *greater than, less than, or equal to* the force on Cart B by Cart A? Explain.

Figure 3.2. Target question 1: the mine cart question.

Specifically, intuitively, it does not seem likely that the motion of the heavy spaceship would be affected by the collision with a light asteroid in a significant way. As such, it is appealing to conclude that the acceleration of the heavy spaceship will be less than that of the asteroid. For most students, the path to the correct answer does not require the rejection of a first available intuitive response. We hypothesized, however, that those students who possess a high level of cognitive reflection skills would be more likely to provide a correct and complete justification for their answers. We argued that these students may have developed the habit of mind to check for consistency between formal reasoning and intuitively appealing responses (even if these responses appear to be highly plausible). As such, target question 2 allowed us to test hypothesis 2: given that students answer a screening question correctly, those students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to recognize the need for justifying their answer to target question 2, even if the correct answer does not require a rejection of the first available intuitive response.

Student responses to all screening-target questions were used to test hypothesis 3: students who possess higher cognitive reflection skills (as measured by the CRT) are more likely to answer the screening-target sequence consistently.

A spaceship is drifting through outer space with its engine off. Suddenly, a small asteroid smashes into the side of the spaceship. (Luckily, the asteroid does not smash *through* the spaceship.)

During the collision, is the magnitude of the acceleration of the spaceship *greater than, less than,* or *equal to* that of the asteroid? **Explain.**

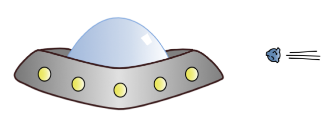


Figure 3.3. Target question 2: *the spaceship.*

### 3.3.3. Data collection

The screening-target questions were included in a single exam. In order to minimize possible biases in student responses (*e.g.,* cuing, priming), these questions were not offered in a sequence [11, 24, 68–70]. Instead, they were distributed throughout the exam. Only about a third of all exam questions were offered in the free-response format, including both target questions. To avoid the placement of the three screening-target questions in close proximity to each other, the screening question was included in a multiple-choice section of the exam. This decision was supported by results of prior research suggesting that correct answers to the screening question are typically accompanied by correct reasoning. At this point in our investigation, it is not clear whether or not the results would be altered if the three research tasks were offered in a sequence. However, by implementing the current research design we attempted to ensure that an entry point into a reasoning path on each question is not immediately influenced by a response on a preceding, analogous task.

In addition to graded work (*e.g.,* exams and homework), students were asked to complete participation-based assignments for assessment purposes either outside of class in a web-based format (*e.g.,* ungraded pre-tests) or in class in a paper-based format (*e.g.,* multiple-choice conceptual evaluations). In both courses, the CRT was included in one of the regular pre-tests. Students received credit for taking the CRT based on their effort as opposed to the correctness of their responses.

### 3.4. Data analysis and results

Below we discuss two techniques used for testing the three hypotheses proposed in the study. First, we established the link between CRT scores and student reasoning approaches by comparing the CRT scores of those students who succeeded on each task and those who did not. Then, we conducted logistic regression analysis in order to determine the relationship between the likelihood (or probabilities) of successful reasoning and CRT scores.

#### 3.4.1. Establishing the link between CRT scores and student reasoning approaches

The average CRT score for the student population in this study is $< CRT >= 1.8$. This number is lower than that reported by Wood *et al* $< CRT_{Wood} >= 2.4$ and somewhat higher than the average CRT score of a general population measured by Frederick ($< CRT_{Frederick} >= 1.24$) [28, 37].

In both semesters, 144 students provided responses to all research tasks used in this study (*i.e.,* the screening-target questions and the CRT). Approximately 85% of these students (N=122) answered the screening question correctly. As stated above, since our goal was to probe the link of cognitive reflection skills on student reasoning approaches in the presence of a mindware, in the following analysis we only included responses from those students who answered the screening question correctly. Specifically, of those students, only ~75% answered target question 1 correctly with correct reasoning. This result indicates that, although the majority of students were able to recognize the applicability of Newton's third law to a challenging screening question, approximately a quarter of these students failed to do the same on a question that elicited a strong intuitive response (*e.g.,* "*$N_{BA}$ is greater than $N_{AB}$ because cart A has much less mass than cart B.*").

Table 3.1. Average CRT scores.

| | Incorrect | Correct | $U$ | $p$ | CL |
|---|---|---|---|---|---|
| Target question 1 ($N = 122$) | $\langle 1.50 \rangle$ | $\langle 1.96 \rangle$ | 1714 | 0.038* | 0.62 |
| Target question 2 ($N = 98$) | $\langle 1.64 \rangle$ | $\langle 2.09 \rangle$ | 1291 | 0.044* | 0.62 |
| | Inconsistent | Consistent | $U$ | $p$ | CL |
| All tasks ($N = 122$) | $\langle 1.54 \rangle$ | $\langle 2.13 \rangle$ | 2425 | 0.002** | 0.65 |

$^{*}p < 0.05$, $^{**}p < 0.01$

The average CRT score of the students who answered target question 1 correctly ($< CRT^{Target1}_{Correct} >= 1.96$) is higher than that of those students who answered incorrectly ($< CRT^{Target1}_{Incorrect} >= 1.50$). Statistical analysis revealed that the difference is significant with a common language effect size (CL) of 0.62 (Mann-Whitney $U = 1714$; two-sided $p = 0.038$; see Table 3.1). According to McGraw and Wong, the common language effect size is the probability that a case randomly selected from one group will have a higher score than a case randomly selected from the other group [71]. A CL of 0.62 is comparable to a Cohen's-$d$ of 0.4 and thus can be considered a 'medium' effect [72, 73]. These results provide evidence in support for hypothesis 1 that, in the absence of a mindware gap, a link exists between the student success on target question 1 and cognitive reflection skills.

Approximately 80% (N=98) of the students who answered the screening question correctly provided correct answers to target question 2, the spaceship task. However, only ~70% of these students included a correct and complete argument in support of their answer. Statistical analysis revealed a significant difference in the distribution of CRT scores of those students who perceived the need to justify their answers and those who did not, as shown in Table 3.1 ($U = 1291$; two-sided $p = 0.044$; CL$= 0.62$). This result provides evidence for hypothesis 2 suggesting that reasoners, who tend to engage in analytical thinking, are more likely to analyze a given problem beyond an intuitively "obvious" answer or a response that "just feels right."

Further analysis revealed that, of those students who answered the screening question correctly, only ~50% gave correct and complete responses to both target questions. The average CRT score for the students who answered all questions consistently (*i.e.,* provided correct answers to all tasks with correct and complete reasoning on target questions) is $< CRT_{cons} >= 2.13$, while the average CRT score for those who provided inconsistent responses is $< CRT_{incons} >$=1.54 ($U$ = 2425; two-sided $p = 0.002$; CL= 0.65). This result provides evidence in support of hypothesis 3.

The results above suggest a link between cognitive reflection skills and student reasoning on questions that tend to illicit strong intuitive responses. Specifically, those students who tend to provide correct, complete, and consistent arguments also tend to receive a higher CRT score. In order to probe this link even further, logistic regression analysis was performed that allowed us to compare the likelihood (or probabilities) of successful reasoning for students with higher and lower cognitive reflection skills.

### 3.4.2. Logistic regression analysis

Logistic regression is used when the dependent (outcome) variable is dichotomous (*i.e.,* 0 or 1). The independent (predictor) variables can be any kind. Logistic regression is utilized in many fields including machine learning, medical fields, and social sciences. Typically, it is used to predict the probability of occurrence of a specific event (*e.g.,* pass/fail, correct/incorrect, survived/perished). The regression algorithm generates the coefficients ($\beta_0, \beta_1, ..., \beta_k$) for each predictive variable $x_k$, which could be used to determine the logarithm of odds that a specific event will occur:

$$log(odds) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k. \tag{3.1}$$

Table 3.2. Logistic Regression coefficients for three models.

| Model | Predictor | Coef. |
|---|---|---|
| Target question 1 ($N = 122$) | Intercept | 0.44 |
| | CRT Score | 0.39* |
| Target question 2 ($N = 98$) | Intercept | −0.02 |
| | CRT Score | 0.42* |
| All Tasks ($N = 122$) | Intercept | −0.92* |
| | CRT Score | 0.53** |

*$p < 0.05$, **$p < 0.01$

In Eq. 3.1, the odds are defined as:

$$odds = \frac{p}{1-p} = \frac{\text{probability of an event occurring}}{\text{probability of an event not occurring}} \tag{3.2}$$

Therefore, the probability for the model with a single predictor, $x$, takes the form:

$$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}. \tag{3.3}$$

In this study, three logistic regression models were generated. In two models, the CRT score was used to predict the probability that each target question was answered correctly. In the third model, the CRT score predicted the probability of consistency in student responses. As shown in Table 3.2, the regression coefficients for the CRT score are statistically significant in all three models, suggesting a trend. However, the interpretation of these logistic regression coefficients is not as straight forward as for the more commonly used ordinary linear regression. For example, according to Eq. 3.1, the logistic regression model for target question 1 suggests that an increase in the CRT score by one unit results in the 0.39 increase in the logarithm of the odds of answering target question 1 correctly. It is, therefore, more illustrative to use logistic regression coefficients either to calculate the odds ratios or to generate graphs of probability functions.

Table 3.3 contains the odds ratios that allow for the comparison between the odds for a successful completion of a task by students with a higher CRT score and the students who scored zero. For example, the odds ratio for students who scored 3 ($Odds_3$) and those who scored 0 ($Odds_0$) suggests that the students with a CRT=3 are ~3.2 times more likely to answer target question 1 correctly, ~3.6 times more likely to recognize the need for justifying their answers on target question 2, and ~5 times more likely to answer all questions consistently. The results suggest that the likelihood of success on each task is related to the tendency of the individual to mediate intuitive thoughts by engaging in a more productive analytic reasoning. Equivalently, this relationship can be visualized by plotting the probability (determined from the odds, $odds = p/1-p$) vs. CRT score for each problem, as shown in Fig. 3.4. All three probability curves follow roughly identical trends, suggesting similar dependency on a CRT score. For example, for the two target questions, the probability of success on each question for the students with a CRT score of 3 is close to 0.8, while the probability of success for those with a CRT score of 0 is between 0.5 and 0.6. The probability for consistency in student responses is also higher for the students with a higher CRT score. The systematically lower probabilities for target question 2 and for consistency reflect the lower student success rate. The insignificant intercept coefficients for the two target questions indicate that the intercept was not significantly different than 50%. The intercept is not important for calculating odds and therefore we are not concerned with its statistical insignificance.

It is important to note that measures of fit and tests of significance for logistic regression are not identical to those used for ordinary linear regression. While in ordinary linear regression,

Table 3.3. Odds ratios comparing the odds of answering each N3L question. Comparisons are made between CRT scores of 1, 2, 3 and CRT=0.

| | Odds Ratio | | |
|---|---|---|---|
| | $\frac{Odds_1}{Odds_0}$ | $\frac{Odds_2}{Odds_0}$ | $\frac{Odds_3}{Odds_0}$ |
| Target question 1 | 1.5 | 2.2 | 3.2 |
| Target question 2 | 1.5 | 2.3 | 3.6 |
| All Tasks | 1.7 | 2.9 | 5.0 |

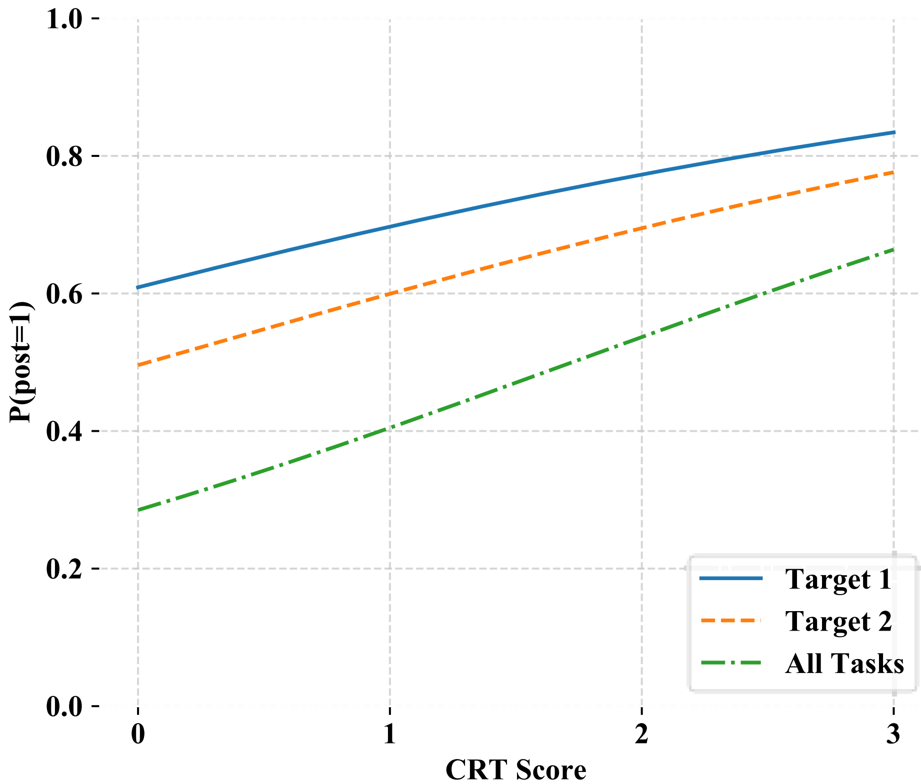



Figure 3.4. Probability curves showing the probability of answering each N3L question. 'All tasks' shows the probability curve of answering consistently.

$R^2$ is used as a single, widely accepted measure of goodness of fit of the model, no common measure of the goodness of fit exists for logistic regression. In our study, two statistical techniques were employed. First, the likelihood ratio $\chi^2$ test for overall model fit was used to compare the accuracy of the null model (not discriminating between different CRT scores) to that of the full model, which includes the CRT score as a predictor of student performance. Results indicate that, for all three cases, the accuracy of the full model is significantly higher than that of the null model (Target 1: $\chi^2 = 4.04$, $p = 0.044$; Target 2: $\chi^2 = 3.95$, $p = 0.047$; All Tasks: $\chi^2 = 9.27$, $p = 0.002$). Second, Receiver Operating Characteristic (ROC) curves were generated for each model in order to estimate the effect size of cognitive reflection on student performance. Each model yielded the area under the ROC curve (AUC) between 0.6 and 0.65 (Target 1: AUC= 0.62; Target 2: AUC= 0.62; All

Tasks: AUC= 0.65). The AUC is equivalent to the $U$-statistic reported in section A, and as such, the effect size calculated through the logistic regression analysis is consistent with the 'medium effect size' obtained from the comparison of CRT averages [74].

### 3.5. Discussion

The results of the data analysis above provide support for a set of hypotheses suggesting that, in the presence of relevant mindware, a positive relationship exists between cognitive reflection skills and student performance on Newton's third law tasks that tend to elicit strong intuitive responses. A screening-target methodology was used to disentangle (to the extent possible) student knowledge and skills from their reasoning approaches. Specifically, student performance on the screening question facilitated identification of those students who possessed necessary mindware to answer the target questions correctly. These students demonstrated that they are able to apply Newton's third law correctly on a task that could be characterized as challenging, but not eliciting strong intuitive responses. A comparison of the average CRT scores of the students who answered target questions 1 and 2 correctly and consistently allowed us to establish the link between cognitive reflection skills and student performance on these tasks. In particular, the average CRT score of the students who answered target question 1 correctly was significantly higher than that of the students who failed to reject an intuitively appealing (but incorrect) response by building a formal argument based on Newton's third law. A similar comparison of the average CRT scores was observed for target question 2 suggesting that the students with higher cognitive reflection skills are more likely to provide formal justification for their answers even if these answers do not require the rejection of the first available intuitive mental model. Finally, those students who answered all three tasks correctly and consistently also exhibited higher cognitive reflection skills (on average) compared to those students who did not succeed in doing so.

The analysis of three logistic regression models allowed for the establishment of a more precise relationship between each CRT score and the likelihood of success on each task. Specifically, the logistic regression model for target question 1 suggests that a one point increase in the CRT score predicts an increases in the odds of a correct answer by a factor of $e^{0.39} \approx 1.5$. Alternatively, this result suggests that the probability of success on target question 1 increases with each increase in the CRT score. This finding provides support for hypothesis 1: in the presence of relevant mindware, on a question eliciting strong intuitive (but incorrect) response, students who possess higher cognitive reflection skills are more likely to mediate intuitive responses by engaging the analytic process more productively. Nearly identical relationships between a CRT score and student success were observed for target question 2, providing support for hypothesis 2: in the presence of relevant mindware, students who possess higher cognitive reflection skills are more likely to recognize the need for justifying their answers, even if the correct response does not require the rejection of the first intuitive mental model. The link between the cognitive reflection skills and consistency in student responses appears to be similar to (if not stronger than) that observed on each of the target questions. The logistic regression model predicts that a one point increase in the CRT score yields an increase in the odds of consistent responses by a factor of $e^{0.53} \approx 1.7$. This result provides evidence for hypothesis 3: students who possess higher cognitive reflection skills are more likely to answer the screening-target questions consistently.

In the hypothesis statements, we stipulated that students must possess mindware. By doing so, we isolated one of the variable that impacts student performance, which allowed us to look for a relationship between reasoning approaches and CRT score. If this stipulation is removed, then we do not find any relationship between CRT and student performance on Target 1, as expected since cognitive reflection alone is not likely to help students arrive at a correct answer. We also have not observed any relationship between student performance on the spaceship question and the CRT performance; this finding is also not particularly surprising because, according to DPToR, in the absence of mindware, CRT is not a reliable predictor of student performance. However, we did observe a relationship between student CRT score and their inclusion of correct reasoning necessary to justify their answer on Target question 2. This result is also not particularly surprising since the criteria imposed on this category of student responses (i.e., correct reasoning) essentially controls for mindware.

While we have shown that a correlation exists between performance on Newton's third law tasks and cognitive reflection skills, the generalizability of these results to other specific contexts remains open for investigation. We speculate, however, that this relationship is not unique to this challenging context. Instead, the results from this study, along with prior findings, suggest that a similar positive relationship may exist between cognitive reflection skills and student performance on many physics tasks that tend to elicit strong intuitive responses. As such, it is plausible that the development of cognitive reflection skills may play a critical role in general scientific thinking and reasoning.

### 3.6. Conclusions

This study was motivated by the growing body of research suggesting that, even after targeted instruction designed to improve conceptual understanding of physics, many students fail to apply the acquired knowledge and skills in a consistent manner. These inconsistencies in student reasoning are particularly persistent on questions that elicit strong intuitive responses. The current study aimed to identify cognitive mechanisms that may be contributing to such inconsistencies. Specifically, according to the Dual Process Theories of Reasoning, a first available mental model (suggested by the fast and subconscious heuristic process) serves as an entry point into a reasoning path. As such, we hypothesized that, given that students acquired relevant knowledge and skills as a result of instruction, students with a higher tendency to mediate first available intuitive responses would be more likely to arrive at a correct answer by engaging the analytic process in a productive manner. The results of this investigation, combined with prior findings, provide evidence for the predicted positive relationship between the cognitive reflection skills and student reasoning in physics. The established relationship is particularly significant because it underlines the importance of designing instructional interventions that attend to both the development of content understanding and strengthening student reasoning skills. While the presence of relevant mindware is necessary for building a correct solution, the ability to systematically reflect on and analyze one's own reasoning path is equally critical for success. We argue, however, that instructional materials that give students explicit opportunities to examine their intuitive ideas are likely to be more effective if offered as a follow up to formal instruction designed to develop solid content understanding. Without the requisite mindware, students would be less likely to recognize the intuitive nature of their responses or would lack formal knowledge necessary for the refinement of their intuitive ideas.

Finally, we speculate that instructional materials and approaches specifically designed to help students understand the interplay between the quick heuristic process and the slow analytic process may provide a twofold benefit to physics instruction by (1) facilitating further improvements in student performance and (2) altering students' attitudes toward physics in general. It is generally known that many students perceive physics to be challenging (See for example [75]). While some students welcome the challenge, for others, physics instruction may generate apprehension and frustration. We speculate that this attitude may stem, at least in part, from a frequent struggle to reconcile the conflict between an answer that "feels right" and formal reasoning based on first principles. The feeling of unease induced by the clash between an appealing (but incorrect) intuitive response and formal reasoning could be illustrated by a statement from a student who answered all screening-target questions correctly and provided the following reflection on his reasoning approaches: “I applied the formal reasoning to the best of my ability, unless I thought myself into a wrong answer, which would look idiotic.” We speculate that physics instruction that makes the dual nature of human cognition explicit and visible to the students may alleviate this struggle to some extent. It may help establish an instructional environment that emphasizes that careful examination (and possible rejection) of an intuitive response is a natural part of the reasoning process and *not* an indication of deficiency or failure. Developing the awareness of one’s own thinking paths as well as the ability to recognize “red flags” is a critical step toward the development of expertise in physics.

# 4. PROBING REASONING APPROACHES AND INTUITION IN THE CONTEXT OF MECHANICAL WAVES

## 4.1. Introduction

In many contexts in physics, students are required to not only apply physics conceptual knowledge in challenging situations, but additionally reason with mathematical approaches. In some situations, significant difficulties occur when a novice must interpret physical meaning in the math. In this study, we aim to explore the utility of dual-process theories of reasoning and the idea of accessibility to interpret student reasoning patterns in situations in which students reason with both physics and mathematics concepts.

Recent research in PER suggests that in contexts in which competing arguments are present and accessible, students will tend to pursue the most accessible option [38, 44]. Therefore, in order to draw from relevant conceptual knowledge, this information must be highly accessible (*i.e.,* easily recalled). We speculate that in general, student applications of intuitive ideas do not stem *solely* from everyday experiences of the real world. Intuition is continuously forming and changing as new information is learned. In particular, like everyday experiences, students may have learned intuitive approaches to problems which involve a mathematical approach and may apply mathematical "shortcuts" in the same way as they apply physical intuition.

The study presented below aims to provide insights into student reasoning approaches in situations which tend to elicit both an intuitive "everyday" approach and an intuitive mathematical application. In the buoyancy case study presented in Chapter 2, we showed that when multiple sources of competing information are at play, some success was found in addressing the everyday intuitive ideas by presenting a physical situation that students could "experience" first hand [38]. Therefore, we aim to explore to what extent the inclusion of highly accessible situations illustrating "real" behavior of mechanical waves impacts student reasoning in situations which may require the application of multivariate expressions.

In particular, we investigate student responses and reasoning approaches concerning wave contexts which have been shown to elicit strong intuitive ideas. Many physical concepts in the context of wave behavior are fundamental for the development of an understanding of advanced physics contexts. The application of the concepts of frequency, propagation speed, wavelength as well as superposition and reflection occur throughout introductory and upper level courses. Typically, these concepts are first introduced through real-world situations such as water waves or waves through a rope or spring. Prior research on student understanding of waves has revealed many difficulties in the contexts of wave pulses, continuous waves at a boundary, interference, and measurements of waves [8, 12, 76–78]. This body of research suggests that these fundamental ideas of wave behavior are very challenging for introductory students. Furthermore, despite considerable efforts into developing effective instructional materials to target these ideas, difficulties in applying fundamental ideas relating to wave behavior persist.

Wittman *et al.* identified several "mental models" which students appear to employ in the given contexts. Specifically, students interpret wave behavior according to a "Particle Pulses" viewpoint, in which wave pulses have a mixture of wave-like and particle-like properties. In this model, students have ideas such as a "harder flick of the wrist" can generate a faster pulse and wave pulses "bounce off" each other. They suggested that many student responses were "dependent on their understanding of the 'trigger' of the question presented."

Further detailed investigations were conducted during the development and refinement of the tutorials on wave pulses and wave interactions at a boundary included in *Tutorials in Introductory Physics*. During these multi-year investigations into student difficulties, student ideas on superposition and reflection of pulses (fixed and free), the interaction of waves at boundaries between different media (*e.g.,* water at different depth) as well as student understanding of the relationship among the concepts of wavelength, propagation speed and frequency were well documented. In particular, students tend to misapply the relationships between the properties of periodic waves often with significant difficulties in recognizing how $\lambda$, $v$, and $f$ can be changed. Further analysis of student responses revealed that many students tend to utilize mathematical 'shortcuts' and manipulate math equations to justify their ideas [8, 12]. Through the lens of DP-ToR, these responses may be interpreted as different cues generating responses from Process 1, in an 'answer first, reasoning second' format [10].

While many student difficulties in these contexts can be attributed to students' misapplication of their everyday experiences, many problems in the context of wave behavior also require the use of multivariate expressions. For many novice students, this may result in a second source of intuitive ideas. In their extensive analysis of how physicists use math, research by Redish and Kuo outlines how the use of mathematical expressions by physicists can appear as if an expert is "filtering the equation through the physics" [79]. Novices in physics, even those with a robust mathematics background, must learn to associate physical meaning with the math. This tendency to have difficulty relating mathematical approaches to physical situations can lead to intuitive mathematical approaches that are not consistent with physical principles [80, 81].

### 4.2. Specific intuitive ideas targeted by the designed interventions

In our previous study in the context of buoyancy, student difficulties stemmed from the competition between relevant and intuitively appealing arguments. We targeted the intuitive ideas which were shown to be present in the buoyancy context and showed that the classroom activities targeting intuitive ideas appeared to give students greater opportunities to examine their first available responses and recognize instances of biased reasoning, leading to higher performance on the posttest [38]. In this study, we aim to apply a similar methodology in our investigation of student approaches to wave contexts in *non-dispersive media*. Specifically, we investigate waves interacting with a boundary and traveling from one medium to another medium (*e.g.,* waves moving from shallow to deep water). In this context, students often treat the propagation speed through one medium as dependent on the speed through another, and will often state that increasing the propagation speed, $v$, for an incident wave will subsequently increase the speed of the wave for the transmitted wave.

Based on the current results from literature, it appears that in this context there are three general reasoning paths that students may take (1) an approach based based on "everyday experiences" (*e.g.,* treat waves as objects)(2) an approach based primarily on treatment of mathematical variables (*e.g.,* plug and chug or symbol manipulation; $v = \lambda/f$, if $f$ increases, $v$ must decrease) (3) an approach based on correct conceptual ideas. As such, students may need to overcome both intuitive everyday reasoning as well as mathematical intuition. We aim to target the intuitive "everyday" idea by including explicit observations of the wave behavior as well as by giving students opportunities to analyze those situations. We speculate that presenting highly relevant examples of the physical situation will make the conceptual ideas more accessible. We hypothesized, therefore, that including simulations and instruction just before students need to apply multivariable expressions may result in reasoning approaches which are more consistent with the physical situation they have just viewed, even if those questions require mathematical applications.

| Possible Reasoning Approaches* | |
|---|---|
| Less Sophisticated | More Sophisticated |
| • Correctly justify observed wave behavior without explicit resolution of intuition about how multivariable expressions work<br>• Incorrectly justify intuitive ideas based on 'everyday' reasoning<br>• Incorrectly justify intuitive ideas based on math 'shortcuts' | • Correctly reflect on intuitive ideas and resolve inconsistencies between `everyday' reasoning and physical principles<br>• Correctly reflect on intuitive math approaches and recognize inconsistencies with the observed wave behavior |

Figure 4.1. Diagram illustrating that correct reasoning may not represent more sophisticated reasoning. *Not an exhaustive list of possibilities.

We recognize, however, that answering the given situation by referencing the prior video does not necessarily represent sophisticated reasoning or deep learning (See Fig 4.1). While we hope that during instruction students will reflect on and resolve any inconsistencies in their own understanding, the *initial* resolution in this case may be no more sophisticated than an intuitive approach. We still view this as positive progress towards building more sophisticated and nuanced understanding.

## 4.3. Methodology

### 4.3.1. Instructional context

This study was conducted during one semester of a large-enrollment introductory calculus-based mechanics course at North Dakota State University. During the semester, many students participate in weekly two-hour lab sessions designed to cover similar topics as those discussed in lecture. Typically, these lab sessions are fairly traditional and do not necessarily closely follow the lecture material. For this study, one of these sessions was replaced by a web-based lab on mechanical waves discussed in detail below. The students worked on the lab in pairs or groups of three during which they discussed their ideas with each other before providing an answer to the questions. In addition, several groups volunteered to have their discussions video recorded for later analysis. The lab was designed for students to progress without significant TA interaction. While a TA was available, a significant guidance from the TA was not expected. During the semester in which the data were collected, mechanical waves were not explicitly covered in detail during lecture, therefore, in addition to the questions developed for this investigation, a brief introduction to waves was also included in the survey.

#### 4.3.2. Overview of the instructional design

One of the goals of this study is to analyze student reasoning approaches (both written responses and dialogue between students' recorded lab sessions) through the lens of DPToR and the theoretical construct of accessibility. In this study, we developed a web-based instructional tutorial in which students investigate the propagation of wave pulses generated on springs (See Fig. 4.3). The instruction was split into roughly four parts: an introduction to wave behavior, a pretest, an intervention and a posttest. We tested the impact of the sequence of questions and accompanying videos based on the following design.

##### 4.3.2.1. Introduction to wave behavior

This section of the lab served primarily to introduce the terminology, symbols, and basic mathematical and causal relationships among wavelength, propagation speed, and frequency. This is not meant to be a comprehensive instruction, but rather an introduction to the ideas that will be utilized throughout the lab. As the students participating in the lab have not necessarily been introduced to the properties of waves in lecture, student performance on this section serves primarily to indicate that students have understood, in some capacity, the definitions and ideas that will be used throughout the lab.

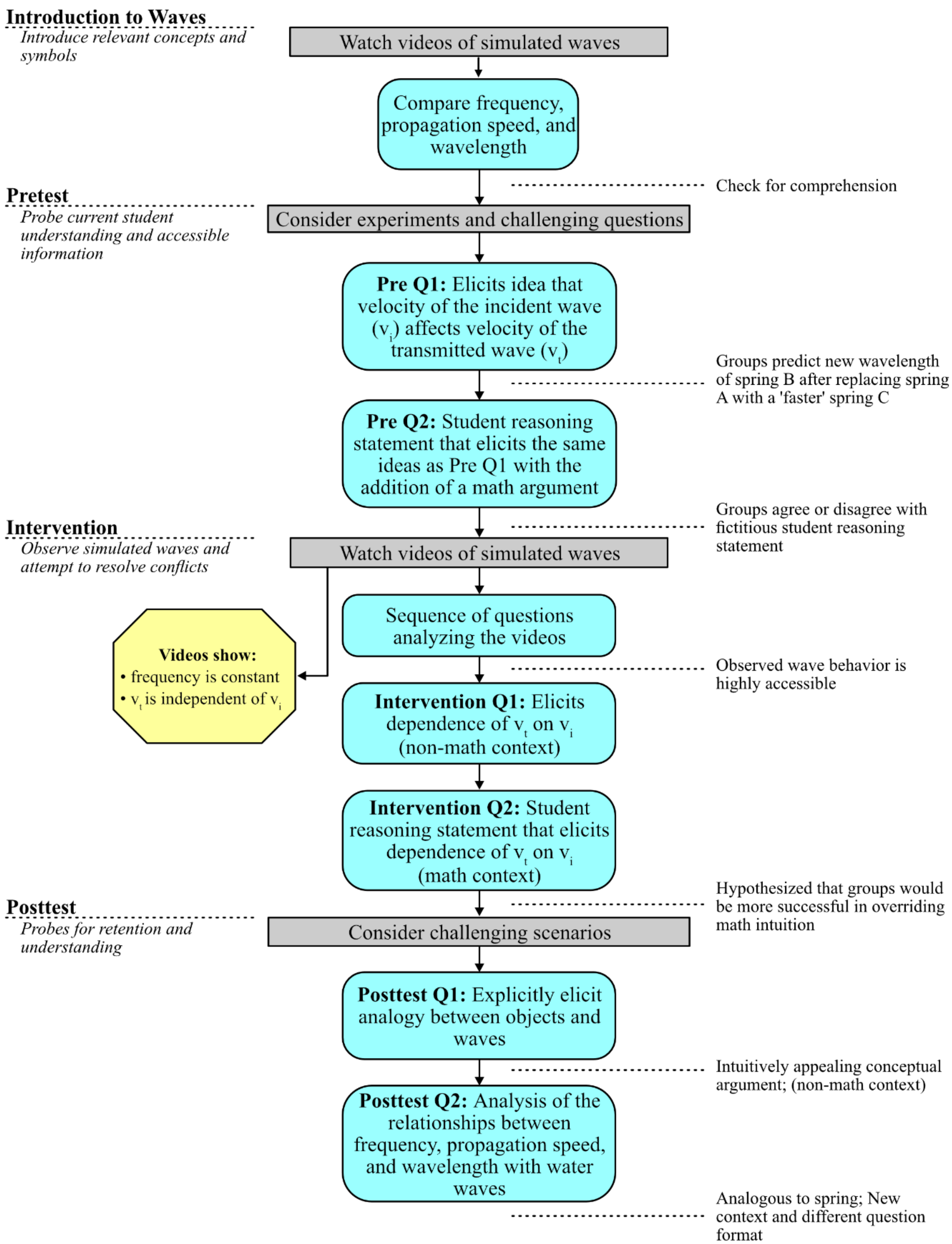



Figure 4.2. Flowchart of the question sequence for the mechanical waves lab.

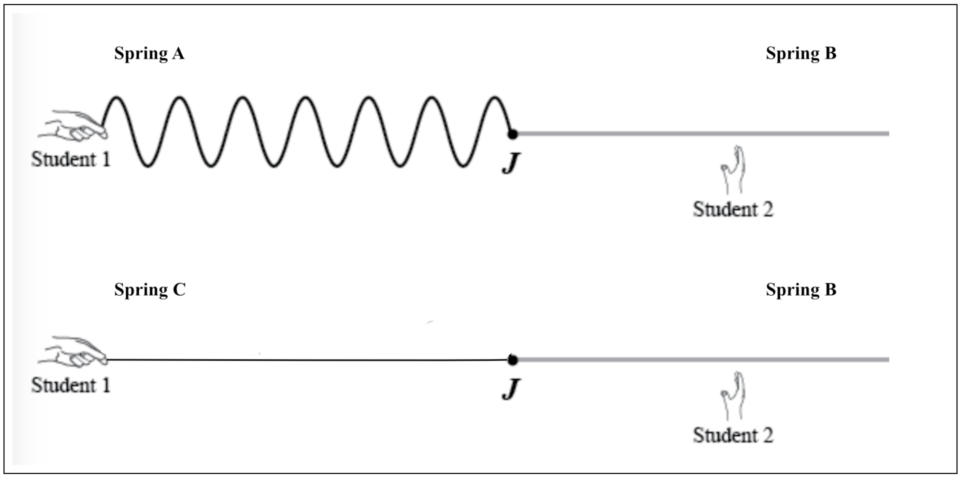



Figure 4.3. Physical situation for PreQ1, PreQ2, and the Intervention.

In this section, students watch videos of simulated springs and answer corresponding questions. The questions have students compare the frequency, propagation speed, and wavelength of the springs shown in the video. If they answer a question incorrectly, they are brought to a page in which explicit feedback is given. The final question asks students to relate their answers to the mathematical expression $\lambda = v/f$ and check for consistency. It is important to note that this is the only section that students are given feedback based on their responses. In the remaining sections, students will attempt to resolve inconsistent ideas on their own.

#### 4.3.2.2. Pretest

The pretest questions probe student current understanding and ability to apply what they learned from the introduction in more challenging situations. Specifically, in the given situations we expect many students will tend to treat the incident waves as particles. In question set 1 (Pre Q1), students consider fictitious students conducting two experiments. As a group, students make predictions about the results of the experiments. In Experiment 1, the speed of a wave in spring A is less than that in spring B. Spring A is then replaced with a different spring (spring C). Students are first guided through questions about the initial experimental setup between spring A and spring B, then are asked to make predictions about how the wavelength for spring B in Experiment 2

In Experiment 2, Spring A was replaced with Spring C, while Spring B remained the same. Student 1 generates a periodic wave with the same frequency as in Experiment 1. Student 1 observes that the waves in spring C travel faster than the waves in spring A in Experiment 1.

Is the wavelength on **spring B** during Experiment 2 *greater than*, *less than* or *equal to* the wavelength on **spring B** during Experiment 1?

Figure 4.4. Question text for Pre Q1.

compares to that in Experiment 1 (See Fig. 4.4). The students are told that waves in spring C travel faster than those in spring A. We use this initial set of questions as a pretest to probe what the lab students are able to do on their own after the basic instruction. In addition, we can examine the reasoning statements in detail and determine what the most common (*i.e., most accessible*) ideas are.

In question set 2 (Pre Q2), students analyze a reasoning statement by a fictitious student who makes an incorrect, but intuitively appealing prediction that "replacing spring A with a 'faster' spring C would increase the speed of a wave in spring B." This line of reasoning is consistent with the common idea that increasing the speed on one side will "transfer" that effect to the other side (*i.e.,* particle pulses). Because students tend to treat waves as objects, we speculate that this situation will elicit strong notions that the student was correct. At this point, the students have not received feedback on the situation so we expect a clear justification of formal ideas to be difficult.

#### 4.3.2.3. Intervention

In this section, students are shown a series of videos of the situation and are asked to answer corresponding questions. These videos build on our previous experience with buoyancy in which explicit activities and references to real-world events appeared to help students recognize that neutral buoyancy is difficult to achieve. Specifically, students are given the opportunity to observe wave behavior, which should help students resolve conflicts between their intuitive ideas and the following ideas: (1) as waves pass from one medium to the other, the frequency stays the same and

Consider the following reasoning statement from Student 1 comparing Experiment 1 and Experiment 2:

*"I know that if the waves travel faster on spring C than on spring A, then the waves must travel faster on spring B as well. We also know that the frequency of the wave remains the same as it travels from spring C to spring B.*

*We can then use the equation λ = v/f to compare the wavelengths.*
*Since the propagation speed is greater in Experiment 2, the wavelength on spring B in Experiment 2 must also be greater than the wavelength on spring B in Experiment 1."*

Do you agree or disagree with the conclusion from Student 1?

Figure 4.5. Question text for Pre Q2.

(2) the propagation speed of incident waves does not affect the propagation speed of transmitted waves. The students then answer a sequence of questions about the videos (*i.e.,* comparing the frequency, wavelength and propagation speed of the waves). By showing the videos immediately after students attempt their own initial predications as well as the analysis of a student response, we aim to make the relevance of the videos more apparent. The full sequence of questions used as well as links to online versions of the videos can be found in the appendix.

The intervention questions, described below, were presented immediately after students watched the corresponding videos. We speculate that because the time between the presentation of the videos and the application of the concept is short, the relevant *conceptual* ideas will be highly accessible. Therefore, we hypothesized that this highly accessible information will compete with any irrelevant information (*e.g.,* mathematical shortcuts) resulting in reasoning approaches which have a greater tendency to rely on conceptual knowledge. The questions investigate (1) what proportion of students recognize that the speed of the incident wave does not affect the speed of the transmitted wave as shown by the simulated waves and (2) after seeing explicitly the physical situation and the interactions between the springs, what proportion of students rely on mathematical shortcuts even in the presence of highly relevant conceptual ideas. For the initial questions, students analyze videos of a simulated wave and describe how the frequency and propagation speed compare in each situation. The students are then asked (Intervention Q1) whether they agree or disagree with the statement *"If the waves travel faster on the spring on the left (incident wave) then the waves must travel faster on the spring on the right (transmitted wave)."* Students are expected to recognize that from the video simulations, the speed of the incident wave does not affect the speed of the transmitted wave.

The last question (Intervention Q2) involved the same experimental situation as Experiment 1 on the pretest, but the second student uses a slightly modified argument. The student in this case also (incorrectly) justified their reasoning using a mathematical expression relating velocity and wavelength (See figure 4.6). The intuitive (incorrect) response in this case is identical to that of Pre Q1, namely that the propagation speed on the left affects the propagation speed of that on the right. We note that at this point in the instruction, students have discussed the situation together in their group and made a prediction (Pre Q1), interpreted and agreed/disagreed with a student statement about the same situation (Pre Q2), and watched a video of a simulation of the situation which shows the result of the experiment. Correct responses to Intervention Q2 would articulate that the conclusion from the student is inconsistent with the video. We hypothesize that student responses to this question will primarily reference the video simulation. In particular, we expect that a high proportion of students will be able to override or ignore any competing mathematical intuition.

#### 4.3.2.4. Posttest

We implemented two posttest question sequences following the intervention. The first question (Posttest Q1) is conceptual in nature (with no reference to the mathematical symbols). In this question, a student makes an analogy between a ball moving from concrete to grass to the behavior of a transmitting wave. Students are expected to articulate that although the analysis

**Recall Experiment 1 and Experiment 2:**

In Experiment 1, it was observed that waves traveled faster on spring B than spring A.

In Experiment 2, spring A was replaced with spring C, while spring B remained the same. Student 1 generated a periodic wave with the same frequency as in Experiment 1. Student 1 observed that the waves in spring C travel faster than the waves in spring A.

Consider the following reasoning statement from Student 2 comparing Experiment 1 and Experiment 2:

*"Since the frequency stays the same, then $v_C/\lambda_C = v_B/\lambda_B$. If the wavelength for spring C is greater than spring A because the speed is greater, then the wavelength for spring B must be greater as well."*

Do you agree or disagree with the conclusion from Student 2? Explain.

Figure 4.6. Question text for InterventionQ2. Accompanying figure corresponds to Fig. 4.3.

is appropriate for objects, waves do not act in the same way. This question is designed to elicit the intuitive "particle pulse" idea. We hypothesize that student ideas are more likely to reference the instruction and videos resulting in reasoning approaches which are more aligned with accepted physical arguments indicating a gain in student understanding.

In the second set of posttest questions, students analyzed propagating water waves in a tank (Posttest Q2, see Fig 4.8). This situation is analogous to the waves in a spring that they have just analyzed, however there are two important differences: (1) the introduction of a new context

Do you agree or disagree with the following statement?

*"I think waves act like a ball rolling from concrete to grass. If we increase the speed of the ball on the concrete, the speed of the ball on the grass will increase proportionally. Similarly, if we increase the speed of the wave on the left, the speed of the wave on the right will increase."*

Explain how this idea is consistent or inconsistent with the simulations you have viewed thus far.

Figure 4.7. Question text for Posttest 1.

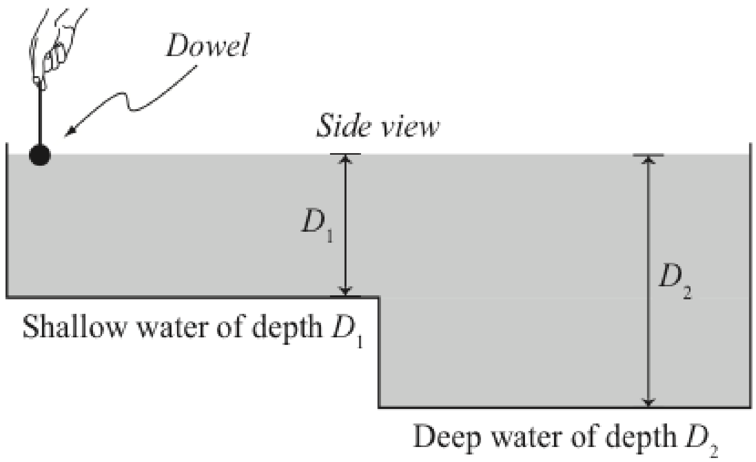



Figure 4.8. Physical situation for PosttestQ2: waves in a tank.

(*i.e.,* students must translate the waves in a spring to water waves) (2) the format of the question has changed from presenting a situation and predicting a result to suggesting a result and having students predict what could be a possible cause. With these differences, student performance on this posttest question is more likely to reflect understanding rather than repetition of answers from the previous instruction. In this case, there are three questions about the waves in a tank scenario, however, the second question in the sequence which asks *"Could changing any of the quantities alone result in the increased wavelength in the right region of the tank?"* most closely probes the concepts we are targeting. Namely, if you change the properties on the left side, does the right side change as well? As such, responses to this question are analyzed in detail as Posttest Q2.

### 4.4. Results

In this study, 56 groups completed the computer based lab representing roughly 130 total students. The students worked in groups and submitted group responses. In addition, 6 groups had their discussion video recorded for analysis. On the introductory information questions, 50% of students correctly answered the first three questions prior to receiving feedback. This indicates those students correctly interpreted the frequency, propagation speed, and wavelength from the

Table 4.1. Summary of student performance on the introduction to wave behavior.

| **Before Feedback** | |
|---|---|
| Comparison | Performance (percentage correct) |
| Frequency | 52% |
| Velocity | 91% |
| Wavelength | 96% |
| All three correct | 49% |
| **After Feedback** | |
| Comparison | Performance (percentage correct) |
| Frequency | 88% |
| Velocity | 82% |
| Wavelength | 98% |
| All three correct | 79% |

video. After initial feedback this proportion increased to 80% of students. We take this as an indication that the majority of students have understood the context and definitions of waves. A summary of the results of this section is presented in Table 4.1.

#### 4.4.1. Pretest results

On the pretest questions, approximately 50% of the groups (N=29) indicated that the frequency should stay the same as the wave pulse propagates from spring A to spring B (See Table 4.2). Similarly, about 50% correctly indicated that since the propagation speed of a wave on spring A is less than that on spring B, the wavelength of the wave on spring A would be less than that on spring B. On the initial predictions of Experiment 2 (Pre Q1), approximately 30% of the groups correctly stated that the wavelength of the wave in spring B would stay the same. 25% of the groups gave the 'particle pulse' answer stating that the increased speed of the incident wave would result in an increased speed of a transmitted wave. 37% of the groups stated that they did not have enough information. Many of the explanations in this case stated that the relationship between the wave speeds on spring B and spring C was not known and so a conclusion could not be formed. Although this information is not relevant to the situation, we feel it is encouraging that the groups were considering alternatives to the intuitive "greater than" answer.

Table 4.2. Summary of student performance on the pretest questions.

| Question | Performance (percentage correct) |
| --- | --- |
| Compare $f$ between spring A and B | 51% |
| Compare $\lambda$ between spring A and B | 50% |
| **PreQ1:** Compare $\lambda$ between spring B (Exp1) and B (Exp2) | 30% |
| **PreQ2:** Agree/Disagree with Student 1 Statement | 52% |

Approximately 50% of the groups (N=29) incorrectly agreed with the student statement (Pre Q2) which predicted that a faster spring on the left (Spring C) would result in an increased speed, and therefore increased wavelength on the right. Predictably, nearly all (90%) of those that correctly answered "equal to" on Pre Q1 disagreed with the statement. Similarly, of the groups that answered "greater than" for their own predictions, over 90% of the groups *agreed* with the statement. Many of the reasoning statements from these groups were similar to the following: *"This is like the last question and we came up with the same conclusion as the student."* This shows that many of these groups followed the same line of reasoning as the student statement, consistent with prior research. Additionally, many groups referenced the equations with reasoning similar to *"It makes sense since the equation shows the relationship."* From these responses, the ideas of "waves as particles" and the tendency to "plug and chug" numbers into multivariate expressions appear to be highly accessible to students.

### 4.4.2. Intervention results

The intervention videos showed that as waves pass from one medium to the other, (1) the frequency stays the same and (2) the propagation speed of incident waves does not affect the propagation speed of transmitted waves. Below we present a typical reasoning path for student groups from one of the recorded video sessions. The question the students are answering is indicated along with minor annotations (*e.g.,* if there was a long pause). In this group the two students, Mark and Barry (pseudonyms), work through the intervention questions, Intervention Q1 and Q2.

*Student Recording:*

[1] **[InterventionQ1]**

[2] Mark: Agree?

[3] Barry: **[repeating question]** If the waves travel faster on the spring on the left, then the waves must travel faster on the spring on the right. I would disagree. So we were wrong before, but I think that was the point of this exercise.

[4] Mark: yeah

[5] **[InterventionQ2]**

[6] Barry: If the wavelength for spring C is greater than spring A because the speed is greater, then the wavelength from spring B must be greater as well. I would disagree with that statement.

[7] Barry: Wait wait. the wavelength for C is greater.

[8] Mark: So, what do we know? We know that f is constant.

[9] Barry: So, frequency stays the same

[10] Mark: That means that lambda A equals $v$ over $f$, which is lambda A has to be less than lambda B. Agreed?

[11] Mark: And we know that lambda A is less than lambda C.

[12] Mark: So what do you?

[13] Barry: If the wavelength is greater, uh, than spring A because the speed is greater, then the wavelength from spring B must be greater as well.

[14] Barry: Isn't that what we just did? Where the [**inaudible**] were all the same?

[15] Mark: If what? What do you mean?

[16] Barry: Isn't that what we just did? In the last, uh...

[17] Mark: Yeah you're right.

[18] Barry: So, it was all the same no matter what.

[19] Mark: So, disagree.

[20] Mark: Since the frequency stays the same. So the faults in the reasoning is...

[21] Barry: Uh, the speed of the propagation is going to be the same. Because of the properties of the spring.

[22] Barry: Right. So spring B will remain [**stopped vocalizing as he was typing**]

We see from the interaction that Barry is reflective of the instruction they have undergone thus far. He seems to recognize that the video simulation was meant to correct their initial responses on previous questions. On the following question, Mark starts an attempt by working out what they "know," and working out the mathematics. Barry again is able to draw from the video experiences and recognize the conceptually similar situation. In this case, they arrive at the correct solution, however neither revisit the mathematics or resolve where the fictitious student went wrong. Instead state the "fault" is that the propagation speed is going to be the same.

Student written responses indicate similar reasoning patterns. After viewing the videos of interacting springs, approximately 85% of the groups gave correct answers to each the questions comparing the frequency, wavelength, and propagation speed of the waves shown in the video simulations. This suggests that although many of the groups were unable to detect the flaw in the initial student statement (Pre Q2), after some instruction nearly all of the groups were able to correctly recognize the correct conceptual ideas. On Intervention Q1, 96% of the groups indicated that they

Table 4.3. Summary of student performance on intervention questions.

| Question Description | Performance (percentage correct) |
|---|---|
| Video analysis question sequence | 85% |
| Intervention Q1 | 96% |
| Intervention Q2 | 70% |

*disagreed* with the statement *"If the waves travel faster on the spring on the left (incident wave) then the waves must travel faster on the spring on the right (transmitted wave)."* This indicates that in a non-math context, most of the groups were able to recognize that the the velocity of the transmitted wave, $v_f$, does not depend on the velocity of the incident wave, $v_i$.

On Intervention Q2, we analyzed those groups that correctly asserted that faster waves on the left do not result in faster waves on the right on Intervention Q1 (See Table 4.3). Approximately 70% of these groups correctly disagreed with the subsequent student statement. Similar to Mark and Barry, many students in this case referenced the video in their reasoning. For example, one group wrote *"We saw from the previous question that the speed and wavelength can be different and produce the same following wave."* This indicates that the relevant information from the video was now highly accessible to many students in analyzing this statement.

Surprisingly, although almost all of the groups correctly analyzed the videos, a third (30%) of those groups did not apply those *same* concepts from the video to the student statement. Instead, 75% of the subsequent incorrect reasoning referred to the equation used with reasoning statements similar to: *"The two springs share a constant $f$. If we solve our equation for $f$ and set them equal to each other we get this formula [$v_C/\lambda_C = v_B/\lambda_B$]."* This suggests that many students may have, upon seeing a correct equation, concluded that the rest of the statement was true as well. In the interaction between Mark and Barry, we see the beginnings of such a reasoning path. Mark's first attempt is to manipulate the mathematics and attempt a solution with the symbols. After drawing attention to the similarity with the previous exercise, both abandon this approach and quickly agree on the answer. This pattern is consistent with those suggested by DPToR. In this case, a first attempt is made using the (possibly) salient mathematical symbols and reasoning from the student statement. However, once a “red flag” appears (explicitly from Barry) a correct solution is quickly produced.

The initial group answers to Pre Q1 serve to indicate prior knowledge of the above situation, however, the responses to Pre Q1 did not predict responses to Intervention Q2. Regardless of the answers to Pre Q1, a similar fraction (approximately 30%) of those who answered ‘equal to’, ‘greater than’, or ‘not enough information’ failed to apply the relevant concepts and instead reasoned with a mathematical approach. Comparing the performance to Pre Q2 suggests that for the questions in which the student dialogue utilized math arguments, the videos had some positive effect (performance increased from 50% to 70% from Pre Q2 to Intervention Q2), however since a large proportion continued to struggle with the multivariable expression, it is unclear whether the videos had an effect on overcoming incorrect mathematical approaches.

We do not claim that the videos brought student groups to more sophisticated reasoning. Indeed, with Mark and Barry we do not see evidence of significant *resolution* of competing ideas. In this case, the correct solution was offered and quickly accepted, but they did not undergo deeper discussion on why Mark’s approach was initially unproductive. It is unclear at this point whether this instruction will have a lasting impact on reasoning and understanding in this context.

#### 4.4.3. Posttest results

On the first posttest question (Posttest Q1), we asked students to consider an analogy between a ball rolling on two different media to that of waves traveling between media (See Fig. 4.7). This situation is challenging as it tends to elicit strong “everyday” reasoning that is difficult for students to overcome. Below we present a segment from the discussion between Phillip and Gary (psudonyms). Prior to this problem, the group was largely successful in reaching a correct conclusion in each of the scenarios presented.

*Student Recording:*

[1] **[Posttest Q1]**

[2] Phillip: **[Reading question]** Do you agree or disagree with the following statement [...]

[3] Phillip: I think no.

[4] Phillip: or well...

[5] Gary: Yeah I mean like he said **[gesturing across the table]**: if it starts faster on the concrete it has to go faster on the grass.

[6] **From across table:** Like in the other case where there were the two [inaudible] and they ended up the same. With the ball.. they don't... I don't know it's not the same...

[7] Gary: Ok so it starts, if this is concrete and this is grass, let's say it goes 10 mph, then here it would slow down right, so like 7. But then if you start at like 40, then it wouldn't go to 7 it would be like 35, or something.

[8] Phillip: I think that kinda depends on the other factors at play here, but it could definitely do that.

[9] Gary: It's just saying if this is larger, then this is larger. It's just saying that if you increase this, then you increase this.

[10] Gary: I think it'd be agree.

[11] Gary: I'm pretty sure it's still talking about this drawing. If they're increasing this speed **[makes wave motion on the left]**, then this speed will increase **[makes wave motion on the right]**.

[12] Phillip: Yeah if they just directly increase it without changing...

On the posttest question, we see that Gary reasoned that the analogy made sense (perhaps influenced from the group across the table). In this case, contrary to their previous answers Gary reasoned that if the left side moves faster, than the right must also move faster. The dialogue between the two students is illustrative of some typical reasoning paths we see in many groups. Students tend to reason that treating waves as particles "make sense," even after viewing the situation explicitly a few minutes prior.

60% of groups correctly disagreed with the statement. When asked how this idea is consistent or inconsistent with the videos they have viewed, many of those groups who disagreed were able to correctly articulate how the videos related to this situation. For example students wrote *"Waves don't act like everyday objects (like the ball) as seen from previous experiments."* and *"This is inconsistent with the simulation of the two waves that started at different speeds, but ended up having the same propagation speed at the end."* Despite many groups articulating how the videos related to the analogy, 40% of groups incorrectly agreed with this statement. Many of these groups persisted with the idea that the speed of the waves in one medium directly affect waves in the right medium. This indicates that despite the nearly identical wording to the questions that relate the to the previous videos (*i.e.,* "if we increase the speed of the wave on the left, the speed of the wave on the right will increase.") the intuitive everyday ideas compete with the ideas from the videos. Indeed, when answering how this idea is consistent with the videos, several groups even appeared to "recall" the consistency. For example, one group wrote, *"The propagation speed of the waves has proportionally transferred after the incident in each example."*

On the next posttest question (Posttest Q2), students consider waves propagating in a tank of water (See Fig. 4.8). This is another challenging situation as it requires students to synthesize their instruction and translate what they learned from the interactions between springs to the context of water waves. Below we present another excerpt from Mark and Barry.

*Student Recording:*

[1] **[Posttest Q2]**

[2] Barry: If we change $v_R$, then yes. If you change the original frequency then yes. But not if you change the original speed. That doesn't affect...

[3] Mark: Waves travel more quickly in deep water. So if we increase this speed, this speed will be increased even more because...

[4] Barry: No, we found that out with the, uh, the wires. It doesn't matter the speed of the original just the speed traveling through the medium.

[5] Mark: But that's different. It's water now because now what we're doing is essentially changing the tension between shallow and deep.

[6] Barry: Ok

[7] Mark: So I'm not sure.

[8] Barry: I dunno. We're also not graded on it so...

[9] Mark: So, what do you think?

[10] Barry: I dunno if we're given some sort of equation like this for water then I don't think it matters of the original speed of these types of waves.

[11] Mark: We're talking about the wavelength. Increase the wavelength of the right region of the tank.

[12] Barry: Mmhmm, I don't think increasing the speed of the first one will transfer to the second one. Just like the wires.

[13] Mark: But it's saying specifically that it does. It says specifically that water waves travel more quickly in deeper water than shallow water. So if we have a velocity in shallow water right? That velocity will increase when it gets to deep. Correct? Which will increase the wavelength. If we increase the original velocity in the lower that would mean we would have an even greater increase in the greater...in the deeper, so the wavelengths should get **[gesturing]** bigger.

[14] Barry: How do we increase the speed of it though?

[15] Mark: It's not wondering whether or not we can. It's saying we changed the velocity... the original velocity. So we know that if $v_1$ is always less than $v_2$. Right? So then if we incr... if $v_1$ goes up then $v_2$ has to also go up. Right. And if the velocity.. and if $v_R$ changes it then $v_L$ has to change it too.

[16] Barry: Ok.

[17] Mark: Does that make sense?

[18] Barry: Sure.

[19] Mark: That's the way I see it. I think.

In this interaction, we see again that Barry is able to draw from what they learned from the waves on a spring, however in this case Mark seems certain that water waves are different. Unlike the situation in Intervention Q2, after drawing attention to the similarity with the "wires," Mark does not abandon the initial approach. This dialogue can emphasize the difficulty in overcoming intuitive ideas in this context. Even with the instruction drawing parallels to waves on a spring and a group-mate explicitly drawing attention to this instruction, the ideas can be difficult to shift.

The written responses show a similar pattern. Looking only at the responses which included $v_L$, we can see which groups indicated that changing the velocity on the left side of the tank would affect the velocity on the right. 60% of groups correctly left $v_L$ unchecked. If we compare this situation to the pretest, on Pre Q1, 30% of students indicated that after replacing spring A with a faster spring C, spring B would remain the same. Although these questions are not identical, the situation with the water waves is challenging and requires thoughtful analysis of each variable. Therefore, a 30% gain is both a statistically significant and instructionally significant increase for the web-based tutorial.

The same Posttest Q2 was given on prior midterm exams. Comparing to those exam scores, after active learning based instruction during regular lecture hours 80% of students correctly left $v_L$ unchecked. There are some significant differences in the instruction that should be considered, however. The lab is a low-stakes environment in which students were working in groups and told that their grade does not depend on correctness. This is in contrast to the high-stakes environment of a mid-term exam. In addition, most student groups spent between 1 - 1.5 hours on the lab tasks, whereas lecture time could consist of 3+ hours between pretest activities, interactive tutorial instruction, and homework. This comparison suggests that although the web-based instruction is not a replacement for in-class active learning lecture, we were still able to produce respectable gains with minimal TA input.

The analysis of the dialogue between students and the analysis of the written responses does not provide substantial evidence that the simulated situation in the videos provided significant impact on the reasoning paths. While a number of groups referenced the videos in the written explanations, a substantial fraction of students relied entirely on a mathematical approach without relating to the video simulation. This may suggest that the situation is similar to that in the buoyancy context presented in Chapter 2. In that case, when instruction emphasized density, the reasoning paths of those who were correct changed from force arguments to density arguments, but no substantial change in overall performance was seen. We speculate that further improvements in performance in the context of wave behavior may require targeted instruction on the mathematical relationships. In particular, targeting intuitive mathematical ideas that are inappropriate or unphysical in the given context.

### 4.5. Discussion

According to DPToR, the heuristic process suggests the first available mental model. Perhaps more importantly, the theories also suggest that this first available response cannot be turned off. Reasoners develop intuition throughout the instructional process and indeed throughout their life. Experts have the advantage that they are able to draw from a large repertoire of experiences and are able to quickly and efficiently "choose" a productive reasoning path. Less experienced students also choose a quick and efficient first reasoning path, however often this initial path is inconsistent with formal accepted reasoning. In this investigation, we probed to what extent making relevant conceptual ideas highly accessible was able to influence reasoning approaches from mathematical intuition. We investigated the intuitive ideas in the context of wave behavior. Prior research has shown that students tend to treat waves and pulses as particles and that students tend to struggle with multivariate expressions. We hypothesized that presenting students with an example of the physical situation would result in explanations which relied on conceptual ideas rather than mathematical approaches. In addition, we believed that presenting the simulated spring interaction and having students answer the questions about the video in groups presents a situation in which students are well primed to tackle additional problems of the same nature. Working in groups affords the individuals with multiple perspectives and more possibilities to check intuitive reasoning. As such, students would have the best chance of reasoning productively after analyzing simulated waves on a spring.

The results show that although over 90% of the participating groups recognized through the video analysis that the reasoning approach shown (*i.e.,* "if you increase the propagation speed of waves incident on the left, the propagation speed of the transmitted waves on the right will also increase") is inconsistent with the video, 30% of those groups did not apply the same reasoning to the analysis of an isomorphic student statement. Instead, the majority relied on a mathematical approach and explanation. We interpret this through the lens of DPToR and accessibility which suggests that the most accessible ideas tend to be considered first. In this case, the instruction provided did not specifically address the application of “simple math.” We speculate therefore that many groups did not feel the need to justify their answer beyond that initial mathematical manipulation.

We note that everyday intuitive ideas are also prevalent in this case. In the case of Phillip and Gary, they struggled to work out the conceptual ideas and relied on more “everyday” reasoning. Mark and Barry also struggled to resolve everyday reasoning on the posttest question in the context of water waves. It is plausible that additional instruction would help many students overcome some of these difficulties, however, as recent research has suggested, instruction which does not target prevalent ideas does not tend to produce significant gains. We speculate therefore that instruction in this context would need to both target the intuitive ‘particle pulse’ ideas and intuitive mathematical ‘shortcuts’ that students tend to use. Currently, curricula have been developed to target the physical ‘particle pulse’ ideas, however it is not clear what effect (if any) additional instruction that includes a discussion of general mathematical arguments would have.

### 4.6. Conclusions

This research was motivated by observed patterns of student difficulties which are resistant to multiple refinements of instruction and continued efforts by highly trained instructors. Inconsistent reasoning and difficulties are particularly persistent on questions and contexts that tend to elicit intuitive responses. We hypothesized that observed patterns of student reasoning could be interpreted through the lens of dual-process theories and that prevalent responses were due to the relative accessibility of ideas. This study aimed to identify accessible ideas in the context of wave behavior and investigate the impact of video simulations of the physical situation on reasoning patterns. We hypothesized that in the given context, viewing the physical situation in a video would help make the relevant conceptual information accessible to students and in turn would help students override difficulties in mathematical reasoning approaches cued by the multivariate expression.

The results of this investigation show that including relevant video simulations appeared to produce moderate gains on conceptual questions, but left a large portion of students struggling with mathematical reasoning approaches. We predicted that student reasoning approaches would reflect the video analysis rather than rely on manipulation of symbols, however, many of the responses (approximately 30%) continued to rely on symbolic arguments. Although most groups could recognize the correct conceptual ideas, applying them to various situations appeared to be difficult. Instead, students tended to use mathematical approaches to the problems that contained multivariate expressions, and in many cases did not apply the ideas from the videos or the questions they had answered. These mathematical approaches and shortcuts appear to be highly accessible and compete with conceptual understanding, even when the relevant physical situation was made apparent.

We argue that this highlights the importance of not only instructing general reasoning skills that apply to physics, but also paying special attention to those contexts in which we expect students to apply particular mathematical approaches. Novice students do not generally check their mathematical answer for physical consistency often reasoning "the equation proves it," referencing an intrinsic trust in the math. Although it is not clear at this point what effect, if any, the video simulations had on the mathematical reasoning approaches, it became clear through the analysis that the mathematical approach was at least as accessible as any other when the symbols were present. We speculate that for future instructional design to be effective, including general mathematical reasoning in physics with particular attention to examples in which the mathematical possibilities are not (or cannot) be realized physically may have the greatest potential to resolve student reasoning difficulties in this context.

In future work, we plan to examine this hypothesis in more detail. In particular, we aim to investigate the impact of including supplemental materials that promote the analysis of mathematical arguments. However, we also emphasize the importance of having activities intended to help students develop conceptual understanding of wave behavior. Many of the reasoning approaches outlined above showed that students need to recognize shortcomings in intuitive ideas both in the concepts and in mathematical approaches. Therefore, it is plausible that a solid understanding of wave behavior may require both mathematical and conceptual approaches to be sufficiently addressed by instruction.

# 5. CONCLUSIONS

Physics education research has identified student difficulties in many contexts and recent efforts have begun to investigate the possibility that reasoning tendencies and biases may play a large role in student response patterns. Recent investigations determined that in order to help students develop productive thinking habits, it is imperative that we direct our efforts towards understanding more precisely the reasoning mechanisms at play in a give context. These results can then be utilized in the design and implementation of refined instructional materials and techniques.

Recent research has additionally shown that the theories from cognitive science on general reasoning tendencies and biases can adequately explain and predict student reasoning patterns on many introductory physics problems. Students do not always answer rationally or consistently, consistent with the idea of 'dysrationalia'. Instead, many patterns of student reasoning seem to stem from an 'answer first, reasoning second' tendency [10]. These patterns are wholly consistent with the ideas of cognitive reflection and accessibility. Knowledge of these heuristic processes have been used to predict the most likely student responses based on processing speed or ease of recall. In addition, it has been shown that students who are more likely to override their initial intuitive ideas are more likely to answer successfully on some types of physics problems. Notably, those with higher cognitive reflection skills, as measured by the CRT, are more successful on common concept inventories like the FMCE and FCI [37, 50].

We have provided further evidence for the ability to explain and predict student tendencies as well as provided initial evidence for systematic ways to improve instruction based on the theoretical framework. In Chapter 2, we investigated whether or not student responses in the context of buoyancy were strongly cued by prominent features of the problem. We first controlled for (to the extent possible) the knowledge and skills acquired from buoyancy instruction. On the exam we gave students a set of questions which tested those concepts that were taught through the tutorial. Then we conducted a series of experiments to investigate the impact of question design and instructional modifications on student reasoning approaches and performance on the "5-block problem." Through this study, we found that in that context, changing question prompts had little effect on performance, but did have a measurable effect on student reasoning approaches to the 5-block problem. In addition, we found that targeting *correct* formal ideas like density of forces did not improve performance. However, the data show that instructional modifications to target highly accessible ideas and student *perceptions* of sinking and floating behavior (*e.g.,* the intuitive notion of "floating in the middle") did significantly improve performance. This study represents and example of strategies that can be used to systematically develop instructional materials based on ideas from reasoning and biases literature.

In Chapter 3 we showed that tendencies in student reasoning, as measured by the CRT, are positively correlated with student performance on some Newton's third law problems. These results are consistent with prior studies that relate CRT scores to FMCE and FCI performance. Additionally, we showed that students with high cognitive reflection skills were more likely to justify their answer with complete physical arguments and answer problems more consistently.

Finally in Chapter 4, we showed how multiple sources of intuitive ideas, not necessarily based on every day intuition, can influence reasoning in difficult contexts. Particularly, for mechanical waves we showed that many students tend to use intuitive mathematical approaches in addition to everyday approaches. While an instructional method to address student intuition based on math approaches was not successfully developed, the results suggests that instruction that targets only one aspect of intuition is likely to be ineffective in shifting student reasoning patterns.

This research has shown that understanding student reasoning patterns and cognitive skills can provide further insights into methods for designing effective instruction. We have shown that it is possible to systematically design instruction based on interpretations of the accessibility of student ideas. We have also shown the importance of reflective skills in performance in physics. Future work in these areas can expand upon each presented study. The buoyancy investigation in Chapter 2 can be extended in several ways. Further refinements of the instructional materials can be made within the buoyancy context to streamline and improve the tutorial additions. In addition, further studies should be conducted to confirm the results of the study in other institutions or in a more diverse set of classrooms. The results of the study can also be extended to other contexts. We suspect that in some contexts student difficulties in overcoming quick, intuitive ideas may be alleviated at least in part with targeted experiences. Similar to those implemented in the buoyancy context, these activities would not target the conceptual ideas, as most lab activities are designed to do, but instead target the most accessible (possibly incorrect) ideas. This design would aim to reduce the plausibility of these incorrect intuitive responses to promote the acceptance and application of formal reasoning approaches.

These same ideas can be applied to the context of mechanical waves to continue to understand reasoning approaches and improve instruction in context in which mathematical intuition plays a significant role. Many contexts, such as those that apply the ideal gas law, $PV = nRT$, contain similar multivariate expressions. We speculate that in these contexts a similar pattern of student reasoning patterns may emerge. Therefore, the results and methods developed in the context of wave behavior could be extended to those contexts as well.

We have shown that cognitive reflection appears to be an important skill for Newton's third law problems. However, there are many contexts in which students answer intuitively in physics. Research which expands the investigations from Newton's third law to other contexts could further investigate the importance of cognitive reflection in physics. In addition, it is not clear what effect (if any) instruction targeting cognitive reflection and intuition may have on reasoning tendencies. Investigations are currently underway in one participating institution to investigate the impact of instruction specifically designed to address and discuss DPToR and intuition. We speculate that the additional instruction that brings the ideas of intuitive responses to the forefront of instruction may have a positive impact on reasoning tendencies in the classroom. Ongoing research projects aim to identify and test instructional approaches to support cognitive reflection in physics classrooms.

I would like to additionally acknowledge Andrew Boudreaux for his contributions to and valuable discussions about these projects. This material is based upon work supported by the National Science Foundation under Grant Nos. DUE-1431857, DUE-1431541, DUE-1431940, DUE-1432765, DUE-1432052, and DRL-0962805.

# REFERENCES


[1] D. Kahneman, *Thinking, Fast and Slow.* Macmillan, Oct. 2011.

[2] J. S. B. T. Evans, “The heuristic-analytic theory of reasoning: Extension and evaluation,” *Psychonomic Bulletin & Review*, vol. 13, pp. 378–395, June 2006.

[3] L. C. McDermott and E. F. Redish, “Resource Letter: PER-1: Physics Education Research,” *American Journal of Physics*, vol. 67, pp. 755–767, Sept. 1999.

[4] L. Hsu, E. Brewe, T. M. Foster, and K. A. Harper, “Resource Letter RPS-1: Research in problem solving,” *American Journal of Physics*, vol. 72, pp. 1147–1156, Sept. 2004.

[5] D. E. Meltzer and R. K. Thornton, “Resource letter ALIP–1: active-learning instruction in physics,” *American journal of physics*, vol. 80, no. 6, pp. 478–496, 2012.

[6] M. E. Loverude, C. H. Kautz, and P. R. L. Heron, “Helping students develop an understanding of Archimedes’ principle. I. Research on student understanding,” *American Journal of Physics*, vol. 71, no. 11, p. 1178, 2003.

[7] P. R. L. Heron, M. E. Loverude, P. S. Shaffer, and L. C. McDermott, “Helping students develop an understanding of Archimedes’ principle. II. Development of research-based instructional materials,” *American Journal of Physics*, vol. 71, pp. 1188–1195, Nov. 2003.

[8] M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, “Student understanding of wave behavior at a boundary: The relationships among wavelength, propagation speed, and frequency,” *American Journal of Physics*, vol. 80, pp. 339–347, Apr. 2012.

[9] L. C. McDermott, “Research on conceptual understanding in mechanics,” *Physics Today*, vol. 37, pp. 24–32, 1984.

[10] M. Kryjevskaia, M. R. Stetzer, and N. Grosz, “Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics,” *Physical Review Special Topics - Physics Education Research*, vol. 10, July 2014.

[11] A. F. Heckler, “The Role of Automatic, Bottom-Up Processes: In the Ubiquitous Patterns of Incorrect Answers to Science Questions,” in *Psychology of Learning and Motivation*, vol. 55, pp. 227–267, Elsevier, 2011.

[12] M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, “Student difficulties measuring distances in terms of wavelength: Lack of basic skills or failure to transfer?,” *Physical Review Special Topics - Physics Education Research*, vol. 9, Feb. 2013.

[13] M. Kryjevskaia, M. R. Stetzer, and T. K. Le, “Failure to Engage: Examining the Impact of Metacognitive Interventions on Persistent Intuitive Reasoning Approaches,” in *Proc. Physics Education Research Conference (PERC 2014)*, pp. 143–146, American Association of Physics Teachers, Apr. 2015.

[14] D. Hammer, A. Elby, R. E. Scherr, and E. F. Redish, “Resources, framing, and transfer,” in *Transfer of learning from a modern multidisciplinary perspective* (J. P. Mestre, ed.), vol. 89, Greenwich, Connecticut: Information Age Publishing, 2005.

[15] I. P. Levin, S. L. Schneider, and G. J. Gaeth, “All frames are not created equal: A typology and critical analysis of framing effects,” *Organizational behavior and human decision processes*, vol. 76, no. 2, pp. 149–188, 1998.

[16] A. Disessa, "Knowledge in pieces," in *Constructivism in the computer age* (G. Forman and P. Pufall, eds.), pp. 49–70, Lawrence Erlbaum, 1988.

[17] L. Louca, A. Elby, D. Hammer, and T. Kagey, "Epistemological Resources: Applying a New Epistemological Framework to Science Instruction," *Educational Psychologist*, vol. 39, pp. 57–68, Mar. 2004.

[18] P. R. L. Heron, "Testing alternative explanations for common responses to conceptual questions: An example in the context of center of mass," *Physical Review Physics Education Research*, vol. 13, June 2017.

[19] L. Lising and A. Elby, "The impact of epistemology on learning: A case study from introductory physics," *American Journal of Physics*, vol. 73, pp. 372–382, Apr. 2005.

[20] M. S. Sabella and G. L. Cochran, "Evidence of Intuitive and Formal Knowledge in Student Responses: Examples from the Context of Dynamics," in *AIP Conference Proceedings*, vol. 720, pp. 89–92, AIP Publishing, Sept. 2004.

[21] R. S. Nickerson, "Confirmation bias: A ubiquitous phenomenon in many guises.," *Review of general psychology*, vol. 2, no. 2, p. 175, 1998.

[22] D. Kahneman, "A perspective on judgment and choice: Mapping bounded rationality," *American Psychologist*, vol. 58, pp. 697–720, Sept. 2003.

[23] D. Kahneman and S. Frederick, "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in *Heuristics and Biases* (T. Gilovich, D. Griffin, and D. Kahneman, eds.), pp. 49–81, Cambridge: Cambridge University Press, 2002.

[24] C. K. Morewedge and D. Kahneman, "Associative processes in intuitive judgment," *Trends in Cognitive Sciences*, vol. 14, pp. 435–440, Oct. 2010.

[25] P. C. Wason, "On the failure to eliminate hypotheses in a conceptual task," *Quarterly journal of experimental psychology*, vol. 12, no. 3, pp. 129–140, 1960.

[26] T. Mussweiler, F. Strack, and T. Pfeiffer, "Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility," *Pers Soc Psychol Bull*, vol. 26, pp. 1142–1150, Nov. 2000.

[27] M. L. Slepian, M. Weisbuch, A. M. Rutchick, L. S. Newman, and N. Ambady, "Shedding light on insight: Priming bright ideas," *Journal of Experimental Social Psychology*, vol. 46, pp. 696–700, July 2010.

[28] S. Frederick, "Cognitive reflection and decision making," *Journal of Economic perspectives*, pp. 25–42, 2005.

[29] J. S. B. T. Evans, "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition," *Annual Review of Psychology*, vol. 59, pp. 255–278, Jan. 2008.

[30] K. E. Stanovich, *What Intelligence Tests Miss: The Psychology of Rational Thought.* Yale University Press, 2009.

[31] S. Ohlsson, *Deep Learning: How the Mind Overrides Experience.* Cambridge: Cambridge University Press, 2011.

[32] R. E. Clark, C. M. Pugh, K. A. Yates, K. Inaba, D. J. Green, and M. E. Sullivan, "The Use of Cognitive Task Analysis to Improve Instructional Descriptions of Procedures," *Journal of Surgical Research*, vol. 173, pp. e37–e42, Mar. 2012.

[33] S. Sloman, "The empirical case for two systems of reasoning," *Psychological Bulletin*, vol. 119, pp. 3–22, January 1996.

[34] V. A. Thompson, J. A. Prowse Turner, and G. Pennycook, "Intuition, reason, and metacognition," *Cognitive Psychology*, vol. 63, pp. 107–140, Nov. 2011.

[35] V. A. Thompson, J. S. B. T. Evans, and J. I. D. Campbell, "Matching bias on the selection task: It's fast and feels good," *Thinking & Reasoning*, vol. 19, pp. 431–452, Sept. 2013.

[36] A. F. Heckler and T. M. Scaife, "Patterns of Response Times and Response Choices to Science Questions: The Influence of Relative Processing Time," *Cogn Sci*, vol. 39, pp. 496–537, Apr. 2015.

[37] A. K. Wood, R. K. Galloway, and J. Hardy, "Can dual processing theory explain physics students' performance on the Force Concept Inventory?," *Phys. Rev. Phys. Educ. Res.*, vol. 12, p. 023101, July 2016.

[38] C. R. Gette, M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, "Probing student reasoning approaches through the lens of dual-process theories: A case study in buoyancy," *Physical Review Physics Education Research*, vol. 14, Mar. 2018.

[39] R. Hertwig, S. M. Herzog, L. J. Schooler, and T. Reimer, "Fluency heuristic: A model of how the mind exploits a by-product of information retrieval.," *Journal of Experimental Psychology: Learning, memory, and cognition*, vol. 34, no. 5, p. 1191, 2008.

[40] L. J. Schooler and R. Hertwig, "How Forgetting Aids Heuristic Inference," *Psychological Review*, vol. 112, pp. 610–628, July 2005.

[41] B. D. Mikula and A. F. Heckler, "Framework and implementation for improving physics essential skills via computer-based practice: Vector math," *Physical Review Physics Education Research*, vol. 13, May 2017.

[42] N. Schwarz, H. Bless, M. Wänke, and P. Winkielman, "Accessibility revisited," in *Foundations of social cognition: A Festschrift in honor of Robert S. Wyer, Jr*, pp. 51–77, Hillsdale, NJ: Lawrence Erlbaum Associates Inc, 2003.

[43] E. T. Higgins, "Knowledge activation: Accessibility, applicability, and salience.," in *Social psychology: Handbook of basic principles.*, pp. 133–168, New York, NY, US: Guilford Press, 1996.

[44] A. F. Heckler and A. M. Bogdan, "Reasoning with alternative explanations in physics: The cognitive accessibility rule," *Physical Review Physics Education Research*, vol. 14, Apr. 2018.

[45] G. Pennycook, J. A. Cheyne, D. J. Koehler, and J. A. Fugelsang, "Is the cognitive reflection test a measure of both reflection and intuition?," *Behav Res*, pp. 1–8, Mar. 2015.

[46] G. Campitelli and P. Gerrans, "Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach," *Mem Cogn*, vol. 42, pp. 434–447, Oct. 2013.

[47] J. M. Liberali, V. F. Reyna, S. Furlan, L. M. Stein, and S. T. Pardo, "Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment," *Journal of behavioral decision making*, vol. 25, no. 4, pp. 361–381, 2012.

[48] M. E. Toplak, R. F. West, and K. E. Stanovich, "The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks," *Memory & Cognition*, vol. 39, pp. 1275–1289, Oct. 2011.

[49] K. E. Stanovich, "Rational and irrational thought: The thinking that IQ tests miss," *Scientific American Mind*, vol. 20, no. 6, pp. 34–39, 2009.

[50] M. Kryjevskaia, C. R. Gette, N. Grosz, M. R. Stetzer, and A. Boudreaux, "Investigating student reasoning with the cognitive reflection test," *Physical Review Physics Education Research*, under review.

[51] L. C. McDermott, P. S. Shaffer, and P. E. G. at the University of Washington, *Tutorials in Introductory Physics*. Upper Saddle River, NJ: Prentice-Hall, 2002.

[52] M. E. Loverude, "A research-based interactive lecture demonstration on sinking and floating," *American Journal of Physics*, vol. 77, pp. 897–901, Oct. 2009.

[53] H. A. Simon, "Invariants of Human Behavior," *Annual Review of Psychology*, vol. 41, no. 1, pp. 1–20, 1990.

[54] R. E. Scherr and A. Elby, "Enabling informed adaptation: Open-source physics worksheets integrated with implementation resources," in *2006 Physics Education Research Conference, AIP Conference Proceedings (PRL Heron, L. Mc-Cullough, and J. Marx, eds.)*, vol. 883, pp. 46–49, 2007.

[55] T. I. Smith and M. C. Wittmann, "Comparing three methods for teaching Newton's third law," *Physical Review Special Topics - Physics Education Research*, vol. 3, Oct. 2007.

[56] S. T. Fiske and S. E. Taylor, "Social cognition, 2nd," *NY: McGraw-Hill*, pp. 16–15, 1991.

[57] P. N. Johnson-Laird, *How We Reason*. Oxford University Press, 2006.

[58] I. A. Halloun and D. Hestenes, "Common sense concepts about motion," *American Journal of Physics*, vol. 53, no. 11, pp. 1056–1065, 1985.

[59] I. A. Halloun and D. Hestenes, "The initial knowledge state of college physics students," *American Journal of Physics*, vol. 53, no. 11, pp. 1043–1055, 1985.

[60] D. E. Brown, "Students' concept of force: the importance of understanding newton's third law," *Physics Education*, vol. 24, no. 6, p. 353, 1989.

[61] D. P. Maloney, "Rule-governed approaches to physics - newton's third law," *Physics Education*, vol. 19, pp. 37–42, January 1984.

[62] D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," *The Physics Teacher*, vol. 30, pp. 141–158, Mar. 1992.

[63] R. K. Thornton and D. R. Sokoloff, "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula," *American Journal of Physics*, vol. 66, pp. 338–352, Apr. 1998.

[64] A. Elby, "Helping physics students learn how to learn," *American Journal of Physics*, vol. 69, pp. S54–S64, July 2001.

[65] D. Hammer and A. Elby, "Tapping Epistemological Resources for Learning Physics," *Journal of the Learning Sciences*, vol. 12, pp. 53–90, Jan. 2003.

[66] A. Maries and C. Singh, "Teaching assistants' performance at identifying common introductory student difficulties in mechanics revealed by the Force Concept Inventory," *Physical Review Physics Education Research*, vol. 12, p. 010131, June 2016.

[67] M. Kryjevskaia, A. Boudreaux, and D. Heins, "Assessing the flexibility of research-based instructional strategies: Implementing tutorials in introductory physics in the lecture environment," *American Journal of Physics*, vol. 82, pp. 238–250, Mar. 2014.

[68] A. A. diSessa, "Toward an Epistemology of Physics," *Cognition and Instruction*, vol. 10, no. 2/3, pp. 105–225, 1993.

[69] E. T. Higgins, "Knowledge activation: Accessibility, applicability, and salience.," in *Social psychology: Handbook of basic principles.*, pp. 133–168, New York, NY, US: Guilford Press, 1996.

[70] E. F. Redish, "A Theoretical Framework for Physics Education Research: Modeling Student Thinking," *arXiv:physics/0411149*, Nov. 2004. arXiv: physics/0411149.

[71] K. O. McGraw and S. Wong, "A common language effect size statistic.," *Psychological bulletin*, vol. 111, no. 2, p. 361, 1992.

[72] J. Cohen, "Statistical power analysis for the behavioral sciences. 2nd," 1988.

[73] W. P. Dunlap, "A program to compute mcgraw and wong's common language effect size indicator," *Behavior Research Methods, Instruments, & Computers*, vol. 31, no. 4, pp. 706–709, 1999.

[74] S. J. Mason and N. E. Graham, "Areas beneath the relative operating characteristics (roc) and relative operating levels (rol) curves: Statistical significance and interpretation," *Quarterly Journal of the Royal Meteorological Society*, vol. 128, no. 584, pp. 2145–2166, 2002.

[75] M. Sobel, "Physics for the non-scientist: A middle way," *Physics Teacher*, vol. 47, pp. 346–349, September 2009.

[76] M. C. Wittmann, R. N. Steinberg, and E. F. Redish, "Making sense of how students make sense of mechanical waves," *The Physics Teacher*, vol. 37, no. 1, pp. 15–21, 1999.

[77] M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, "Student understanding of wave behavior at a boundary: The limiting case of reflection at fixed and free ends," *American Journal of Physics*, vol. 79, pp. 508–516, May 2011.

[78] M. Kryjevskaia, M. R. Stetzer, and P. R. L. Heron, "Is a Simple Measurement Task a Roadblock to Student Understanding of Wave Phenomena?," *The Physics Teacher*, vol. 51, pp. 560–563, Dec. 2013.

[79] E. F. Redish and E. Kuo, "Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology," *Science & Education*, vol. 24, pp. 561–590, July 2015.

[80] E. Kuo, M. M. Hull, A. Gupta, and A. Elby, "How students blend conceptual and formal mathematical reasoning in solving physics problems: Conceptual and Formal Mathematical Reasoning," *Science Education*, vol. 97, no. 1.

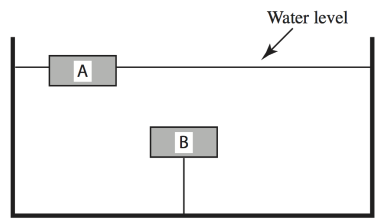
[81] T. J. Bing and E. F. Redish, "Analyzing problem solving using math in physics: Epistemological framing via warrants," *Physical Review Special Topics - Physics Education Research*, vol. 5, Dec. 2009.

# APPENDIX

Two identical wood blocks, A and B, are placed in a tank of water. Block A floats at the surface as shown, while block B is held down with a string.

1. Draw a free-body diagram for block A. Make sure the lengths of your force vectors indicate the correct relative magnitudes. Explain briefly.

2. Is the magnitude of the buoyant force on block A *greater than, less than,* or *equal to* the magnitude of the buoyant force on block B? Explain.




Consider now blocks B and C at right. Block C is held up with a string, as shown. Both blocks are of the same size and shape, but block C is made from metal. The mass of C is greater than that of B, $m_C > m_B$. Without the string, block C would sink to the bottom.

3. Is the magnitude of the buoyant force on block B *greater than, less than,* or *equal to* the magnitude of the buoyant force on block C? Explain.

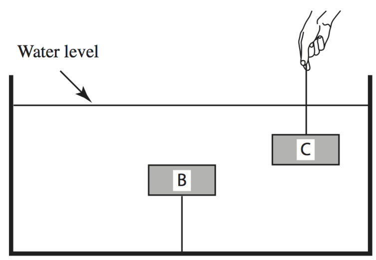



Figure A.1. Sequence of screening questions.

## A.1. Original buoyancy tutorial

In Section I, students consider the behavior of a block submerged in water. Students are told that after the block is released, it is observed to float on the surface of the water. Students are asked: (1) to draw an extended free-body diagram of the block just after it is released in the middle of the beaker of water and to identify all forces, (2) to use the relationship between pressure and depth to compare the magnitudes of the vertical forces, (3) to determine the vector sum of the forces exerted on the block by the surrounding water and to compare this force to the weight of the block. In the following section, students are told that the experiment is repeated with a second block of the same volume and shape; however, the second block is observed to sink. Students are then guided through the same set of questions. Upon completion of this section, the buoyant force is introduced as the vector sum of the forces exerted on an object by the surrounding liquid.

Students are asked to articulate explicitly whether or not the buoyant force on an object that is completely submerged in an incompressible liquid depends on: the mass or weight of the object, the depth below the surface at which the object is located, and the volume of the object. Students are expected to recognize that neither the mass nor the depth of the object impacts the magnitude of the buoyant force.

In Section 2, Archimedes' principle is introduced along with the idea of the volume of liquid displaced by the object.

Section 3 explicitly focuses on sinking and floating. First, students consider block A, which is observed to float on the surface of the water after it had been released from the center of a beaker. Students compare the buoyant force and the weight at two instances: right after block A is released and when the block reaches the surface. Then, students consider a different block, block B, of the same size and shape as block A, but with a slightly greater mass. The block is observed to barely float on the surface of the water. Students compare the buoyant forces on blocks A and B right after they are released and at their final positions. Finally, students consider block C of the same size and shape as blocks A and B, but with a mass slightly greater than that of block B. Students are asked to consider a hypothetical incorrect prediction of the final position of the block given by two fictitious students and to identify incorrect elements of reasoning given by both students. Specifically, Student 1 incorrectly uses the mass of the block to predict that block C will float right below the surface since block C is heavier than block B. Student 2 agrees and supports this incorrect prediction with a force argument: since the buoyant force on block C is slightly less than that on block B, block C should come to rest a bit below the surface. Students are expected to recognize that student 2 ranked the buoyant forces on Blocks B and C correctly, but failed to recognize that the net force on Block C points toward the bottom of the beaker.

## A.2. Screening questions

The force argument that leads to the correct answer to the 5-block problem requires the following logical steps: (1) recognize that the buoyant force is determined by the displaced volume of the liquid (and the density of that liquid) rather than by the object's mass or depth, and (2) draw a free-body diagram and use Newton's Laws to predict the direction of the net force on the object of interest. More specifically, students must recognize that the buoyant force on all the submerged blocks is the same, while the weight increases according to the mass; since block 2 barely floats, the buoyant force on the submerged blocks is approximately equal to the weight of block 2. This suggests that the net force on block 1 is toward the surface of the liquid and block 1 will therefore float; however, the net forces on blocks 3-5 are toward the bottom of the container, so those blocks will sink. The screening questions, shown in Fig. A.1, require students to demonstrate the same knowledge and conceptual understanding outlined above. In the first two questions, students are asked to consider blocks A and B. They are told that both blocks have identical volume and mass. The first question asks students to draw a free-body diagram for block A, thereby identifying both forces acting on the block ($mg$ and $F_{buoyant}$). The following question asks students to compare the magnitude of the buoyant force on block A to that on block B. This question probes whether or not students recognize that, although the blocks are identical, the volume displaced by block A is less than that displaced by block B. Therefore, the magnitude of the buoyant force on A is less than that on B. The final question asks students to compare the buoyant forces on blocks B and C of equal volume, but different mass ($m_C > m_B$). The blocks are completely submerged and suspended at different depths by two strings, as shown in Fig. A.1. Students are told that in the absence of the strings, block B would float, while block C would sink. Once again, students are expected to recognize that only the displaced volume determines the ranking of the buoyant forces in this situation, and therefore the buoyant forces on blocks B and C are equal in magnitude.

## A.3. Mechanical waves lab

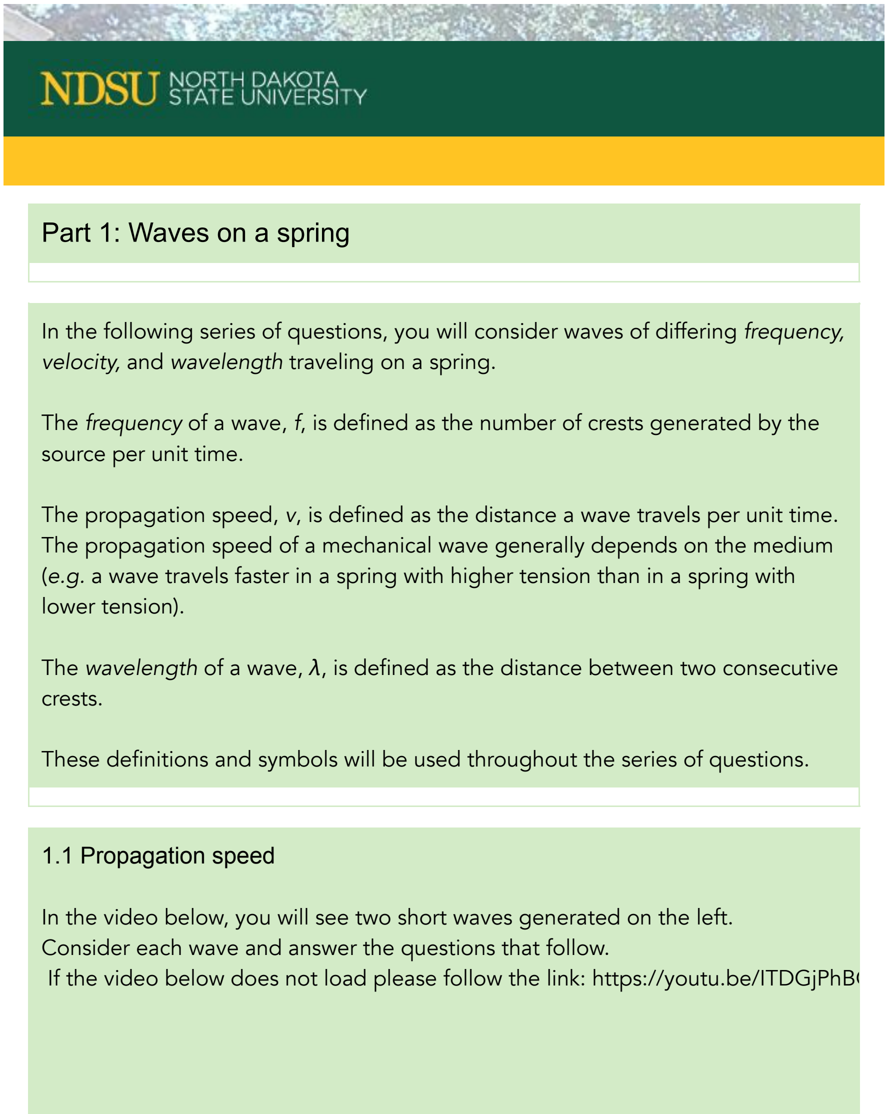



Figure A.2. Online mechanical waves lab.

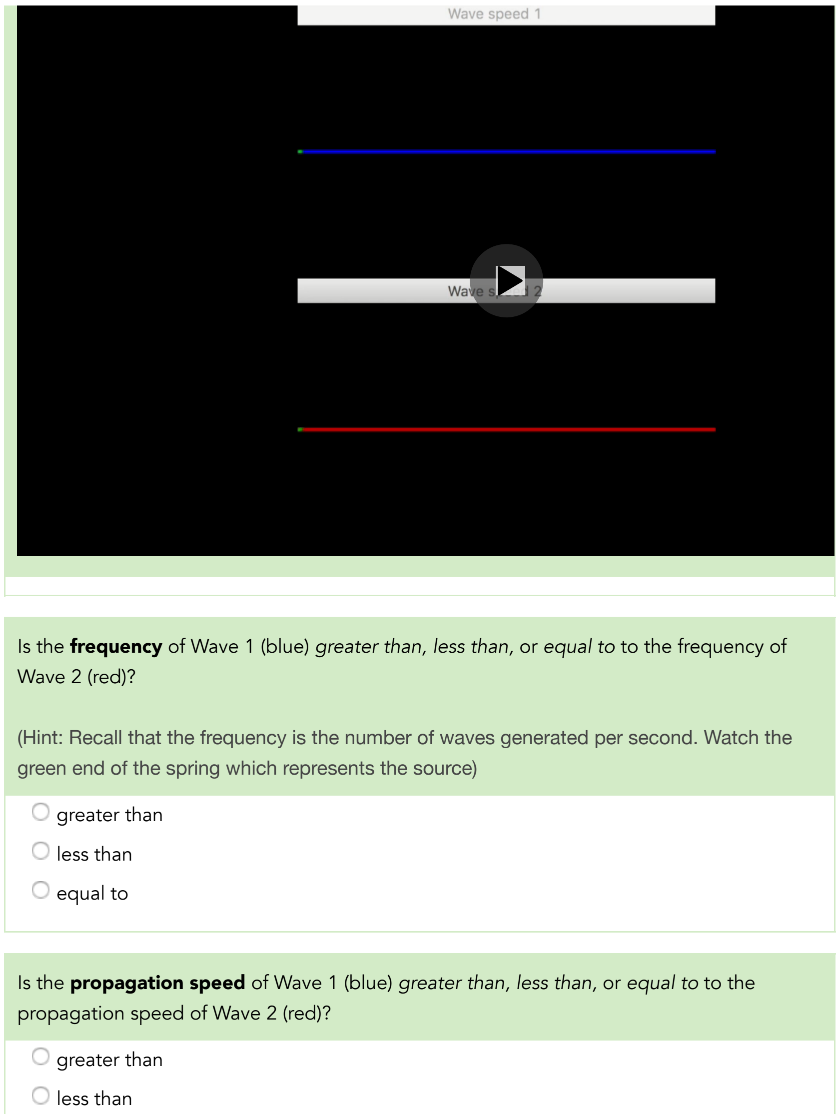
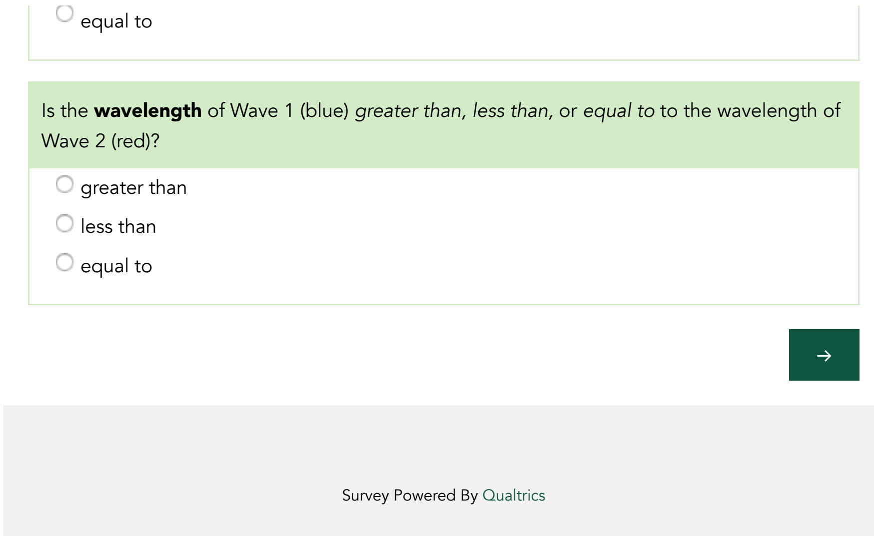



Figure A.2. Online mechanical waves lab (continued).


○ equal to

Is the **wavelength** of Wave 1 (blue) *greater than*, *less than*, or *equal to* to the wavelength of Wave 2 (red)?

○ greater than
○ less than
○ equal to

→

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).

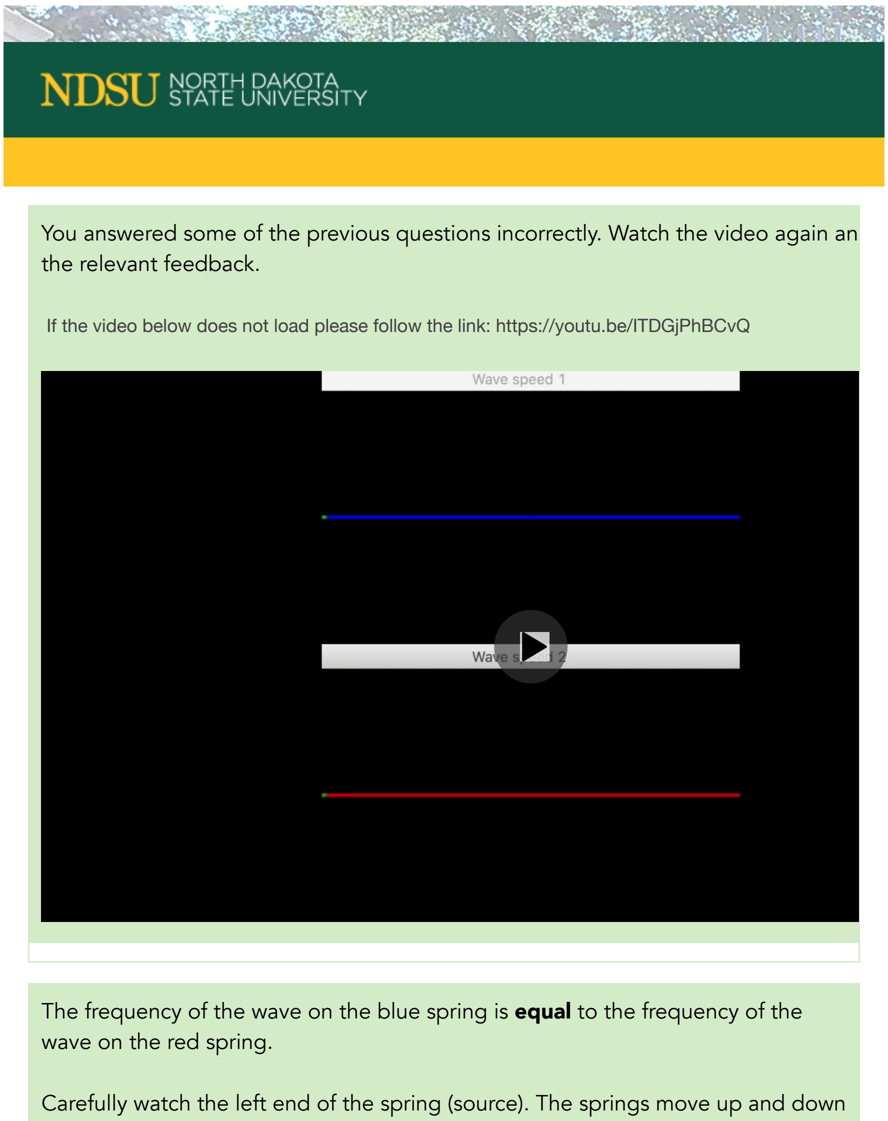



Figure A.2. Online mechanical waves lab (continued).

together, and so the source generates two crests in the same amount of time.

The propagation speed of the wave on the blue spring is **greater than** the propagation speed of the wave on the red spring.

The wavefront (rightmost edge of the wave) in the blue spring moves faster than the wavefront in the red spring.

The wavelength of the wave on the blue spring is **greater than** the wavelength of the wave on the red spring. The distance between the two crests is greater on the blue spring than the red spring.

→

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).

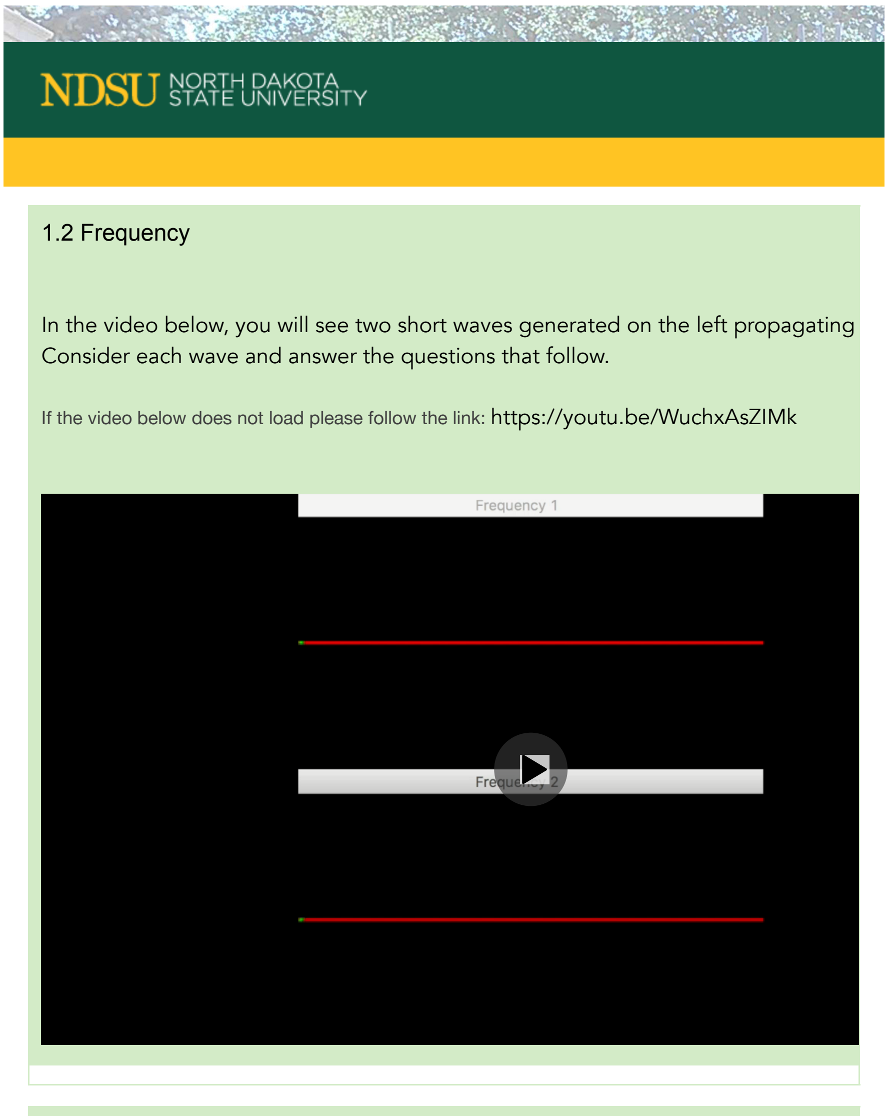



Figure A.2. Online mechanical waves lab (continued).

Is the **frequency** of Wave 1 (top) *greater than*, *less than*, or *equal to* to the frequency of Wave 2 (bottom)?

- ○ greater than
- ○ less than
- ○ equal to

Is the **propagation speed** of Wave 1 (top) *greater than*, *less than*, or *equal to* to the propagation speed of Wave 2 (bottom)?

- ○ greater than
- ○ less than
- ○ equal to

Is the **wavelength** of Wave 1 (top) *greater than*, *less than*, or *equal to* to the wavelength of Wave 2 (bottom)?

- ○ greater than
- ○ less than
- ○ equal to

Analogous to the kinematic relation, $d = v\Delta t$, which relates the distance traveled to the speed and the time, the wavelength, $\lambda$, is related to the propagation speed, $v$, and the time it takes the source to generate a wave, $T$, (also known as the period) by $\lambda = vT$.

Recall that the frequency of a wave is the number of waves generated per unit time. The frequency therefore is related to the period of a wave by $f = 1/T$, which represents that only one wave is generated during one period.

We can combine the two expressions to relate the wavelength $\lambda$, the propagation speed, $v$, and the frequency $f$:

Figure A.2. Online mechanical waves lab (continued).

$$\lambda = \frac{v}{f}$$

Explain how your answers are consistent or inconsistent with the relation above.

In general for a mechanical wave, the propagation speed, *v*, is determined by the medium through which the wave is propagating. For a spring, the propagation speed is determined by the tension, $F_T$, and the linear mass density, μ, through the following relation:

$$v = \sqrt{\frac{F_T}{\mu}}$$

The frequency, *f*, is determined by the source that is generating the wave.

Both quantites, the propagation speed and the frequency, together determine the wavelength, $\lambda$.

Figure A.2. Online mechanical waves lab (continued).

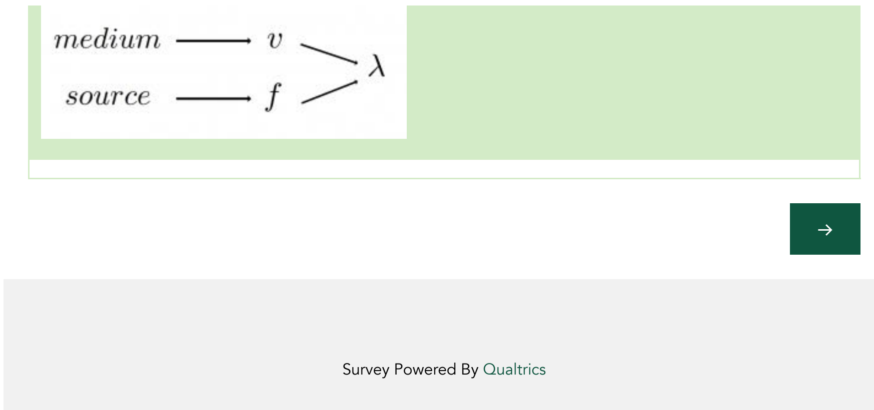


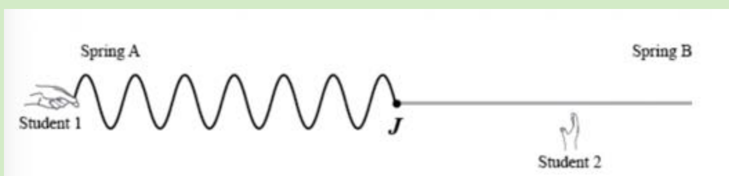

Figure A.2. Online mechanical waves lab (continued).

NDSU NORTH DAKOTA STATE UNIVERSITY

## Part 2: Waves at a boundary

Two springs, A and B, are attached end-to-end at point J as shown below. Student 1 generates a periodic wave on spring A. The figure shows the wave on spring A at one instant; the wave on spring B is not shown.

Student 2, standing near spring B, notes that a wave travels *faster* on spring B than on spring A.




Suppose in Experiment 1, Student 1 measures the wavelength on spring A to be $\lambda_A$ and Student 2 measures the wavelength on spring B to be $\lambda_B$ .

As the wave travels to spring B, does the frequency *increase, decrease*, or *remain the same*?

- ○ increase
- ○ decrease
- ○ remains the same
- ○

Figure A.2. Online mechanical waves lab (continued).

not enough information

Explain your reasoning.

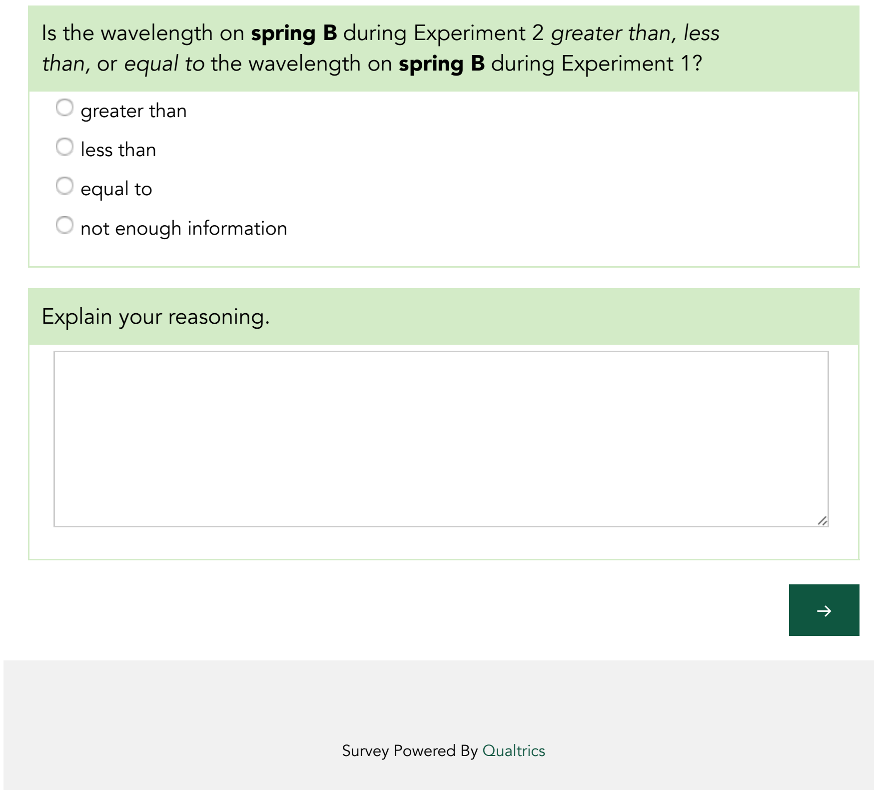
How does the wavelength of spring A compare to that of spring B?

- $\lambda_A < \lambda_B$
- $\lambda_A > \lambda_B$
- $\lambda_A = \lambda_B$
- not enough information

Explain your reasoning.

In Experiment 2, Spring A was replaced with Spring C, while Spring B remained the same. Student 1 generates a periodic wave with the same frequency as in Experiment 1. Student 1 observes that the waves in spring C travel *faster* than the waves in spring A in Experiment 1.

Figure A.2. Online mechanical waves lab (continued).



Is the wavelength on **spring B** during Experiment 2 *greater than, less than*, or *equal to* the wavelength on **spring B** during Experiment 1?

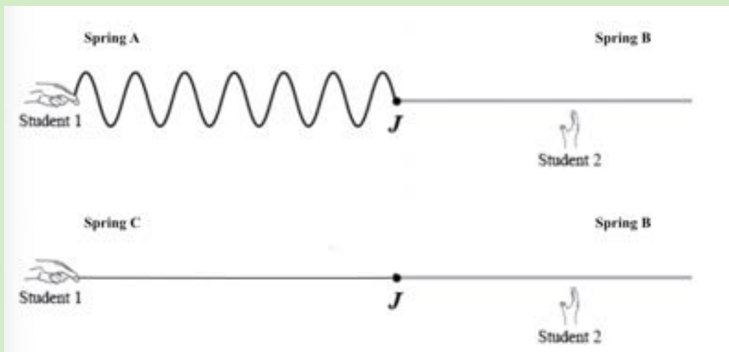
- greater than
- less than
- equal to
- not enough information

Explain your reasoning.

→

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).




Consider the following reasoning statement from Student 1 comparing Experiment 1 and Experiment 2:

*"I know that if the waves travel faster on spring C than on spring A, then the waves must travel faster on spring B as well. We also know that the frequency of the wave remains the same as it travels from spring C to spring B.*

*We can then use the equation $\lambda = v/f$ to compare the wavelengths. Since the propagation speed is greater in Experiment 2, the wavelength on spring B in Experiment 2 must also be greater than the wavelength on spring B in Experiment 1."*

Do you agree or disagree with the conclusion from Student 1?

○ Agree

Figure A.2. Online mechanical waves lab (continued).

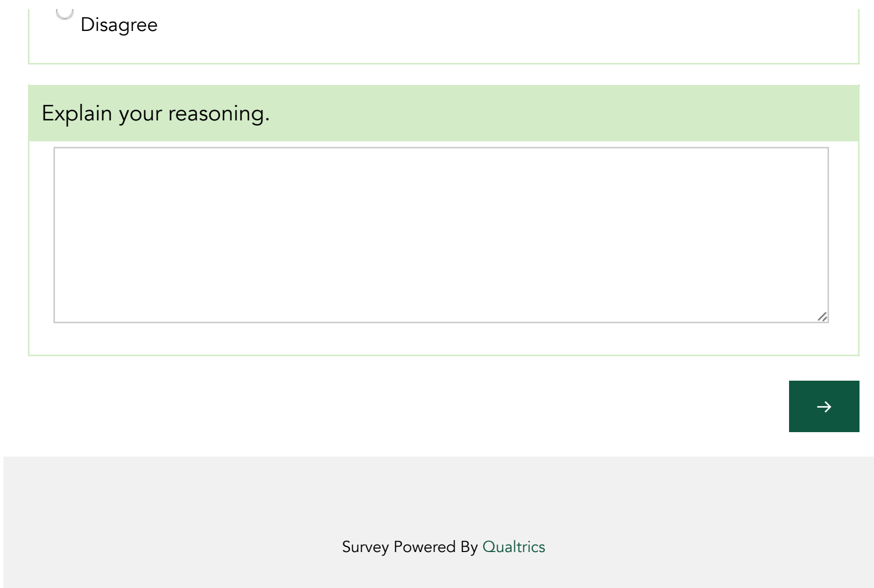
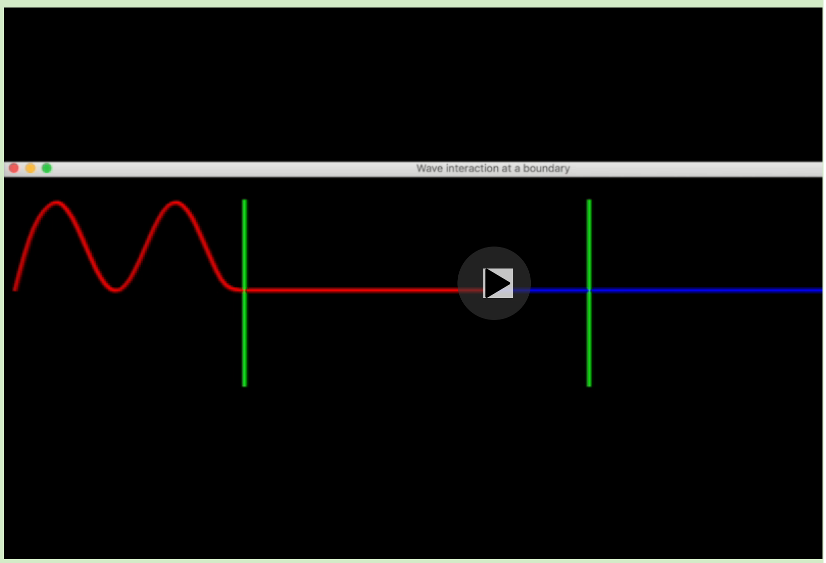



Figure A.2. Online mechanical waves lab (continued).

NDSU NORTH DAKOTA STATE UNIVERSITY

In the video below, you will see a short wave generated on the left propagating to the right.
The wave will interact with a boundary between two springs.
Consider the video and answer the questions that follow.

If the video below does not load please follow the link: https://youtu.be/E9jfUSHj9O0




Is the **frequency** of Wave 1 (red) greater than, less than, or equal to to the frequency of Wave 2 (blue)? (Hint: the green bars flash when the peak passes)

Figure A.2. Online mechanical waves lab (continued).

greater than

- less than
- equal to

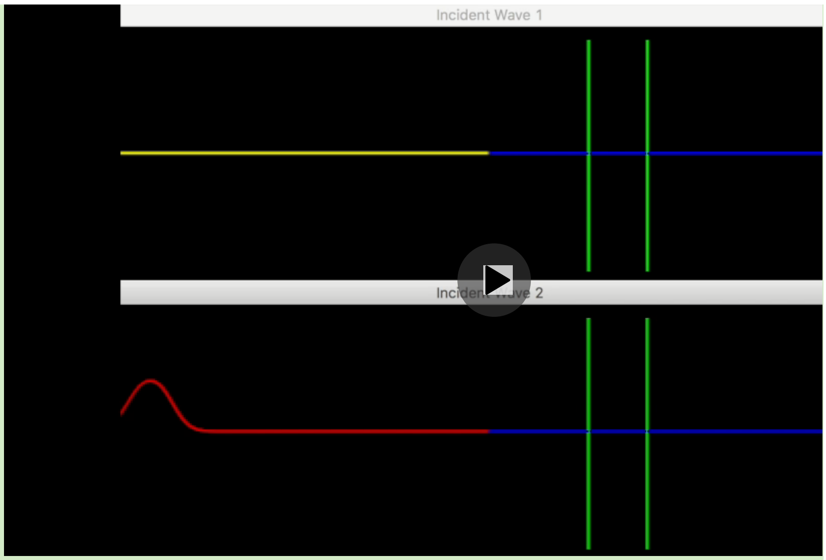
Explain your reasoning.

In the video below, you will see two short waves generated on the left propagating to the righ
Each wave will interact with a boundary between two springs.
Take note of any differences you see and answer the questions that follow.

You can assume that the two incident waves (yellow and red) were generated using the same s
have the same frequency).

If the video below does not load please follow the link: https://youtu.be/o_mWOuC

Figure A.2. Online mechanical waves lab (continued).




Is the **propagation speed** of Incident Wave 1 (yellow) greater than, less than, or equal to to the propagation speed of Incident Wave 2 (red)?

- ○ greater than
- ○ less than
- ○ equal to

Figure A.2. Online mechanical waves lab (continued).

Explain your reasoning.

Is the **propagation speed** of Transmitted Wave 1 (top blue) greater than, less than, or equal to to the propagation speed of Transmitted Wave 2 (bottom blue)?

- ○ greater than
- ○ less than
- ○ equal to

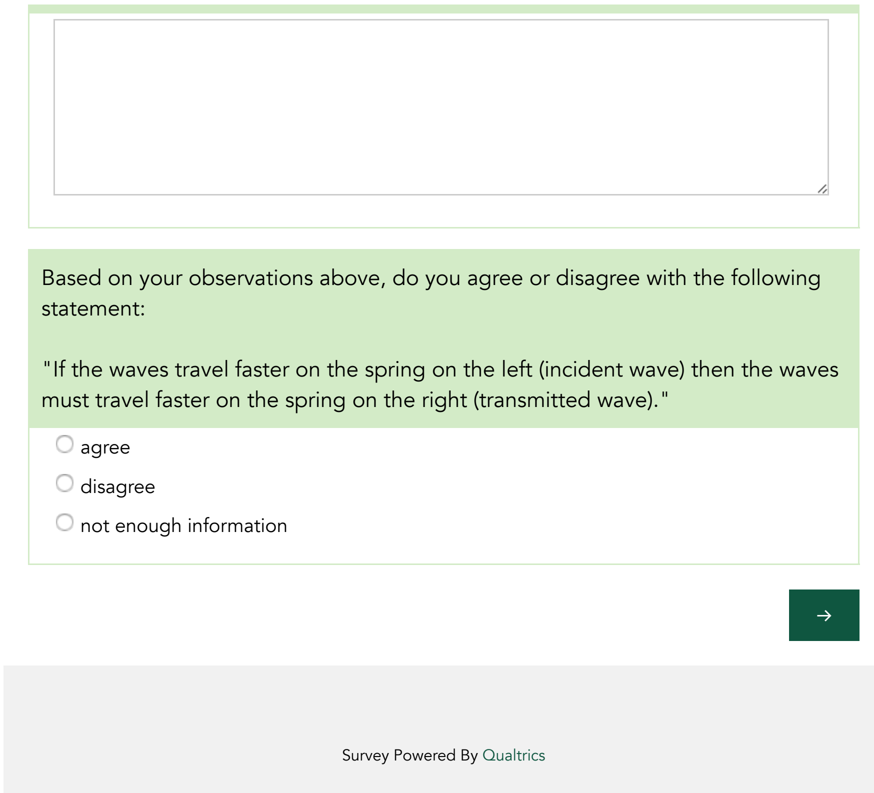
Explain your reasoning.

Is the **wavelength** of Transmitted Wave 1 (top blue) greater than, less than, or equal to to the wavelength of Transmitted Wave 2 (bottom blue)?

- ○ greater than
- ○ less than
- ○ equal to

Explain your reasoning.

Figure A.2. Online mechanical waves lab (continued).


Based on your observations above, do you agree or disagree with the following statement:

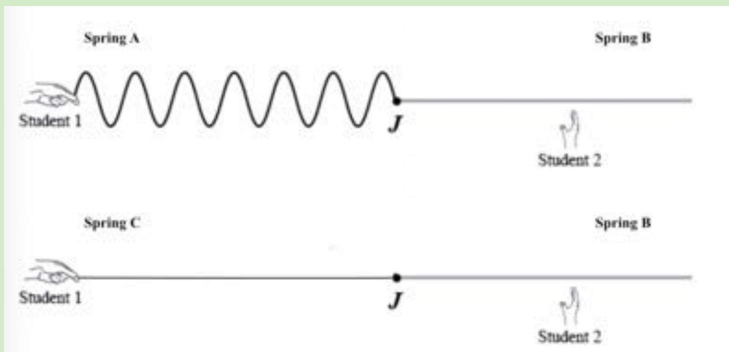
"If the waves travel faster on the spring on the left (incident wave) then the waves must travel faster on the spring on the right (transmitted wave)."

- agree
- disagree
- not enough information

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).

NDSU NORTH DAKOTA STATE UNIVERSITY




Recall Experiment 1 and Experiment 2:

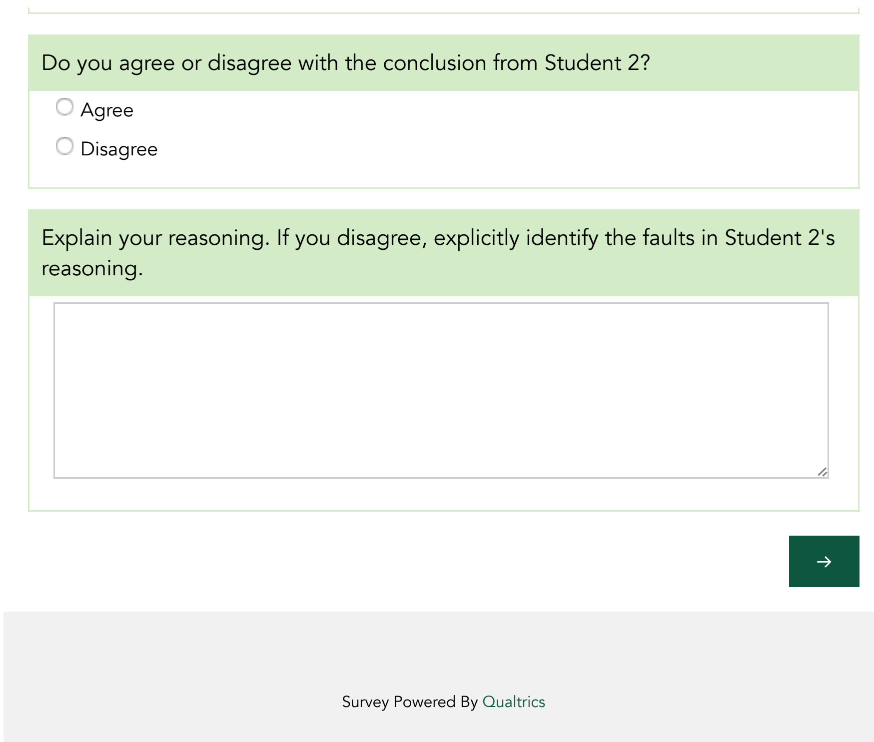
*In Experiment 1, it was observed that waves traveled faster on Spring B than Spring A.*

*In Experiment 2, Spring A was replaced with Spring C, while Spring B remained the same. Student 1 generated a periodic wave with the same frequency as in Experiment 1. Student 1 observed that the waves in spring C travel faster than the waves in spring A.*

***Consider the following reasoning statement from Student 2 comparing Experiment 1 and Experiment 2:***

*"Since the frequency stays the same, then $v_C/\lambda_C = v_B/\lambda_B$. If the wavelength for spring C is greater than spring A because the speed is greater, then the wavelength for spring B must be greater as well."*

Figure A.2. Online mechanical waves lab (continued).


Do you agree or disagree with the conclusion from Student 2?

- Agree
- Disagree

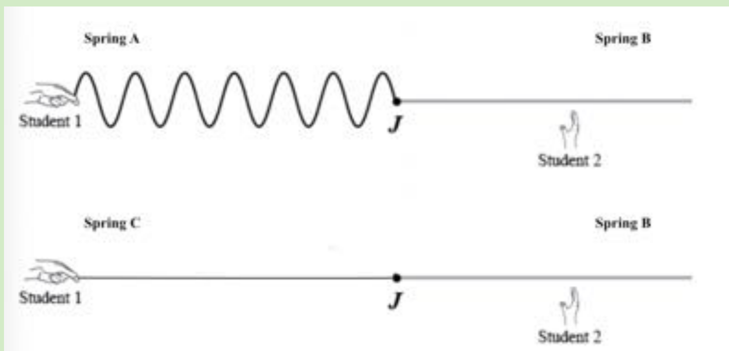
Explain your reasoning. If you disagree, explicitly identify the faults in Student 2's reasoning.

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).

NDSU NORTH DAKOTA STATE UNIVERSITY

Part 3: Analyze reasoning and predictions




Consider the following idea from Student 1 for a new experiment:

"Propagation speed depends on frequency according to the expression $v = \lambda f$, therefore if I double the frequency of the wave in spring C then the propagation speed of the wave in spring C also doubles."

Do you agree or disagree with Student 1's prediction?

- ◯ Agree
- ◯ Disagree

Explain your reasoning. If you disagree, explicitly identify the faults in Student 1's

Figure A.2. Online mechanical waves lab (continued).

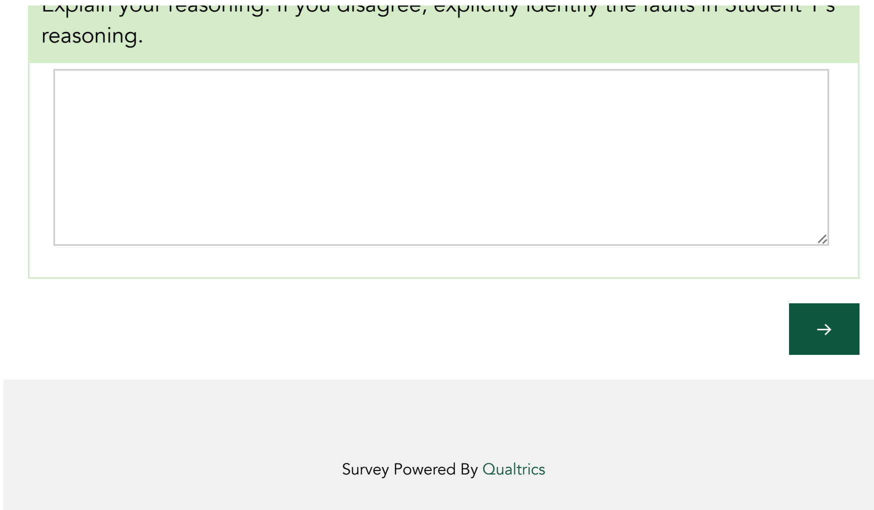
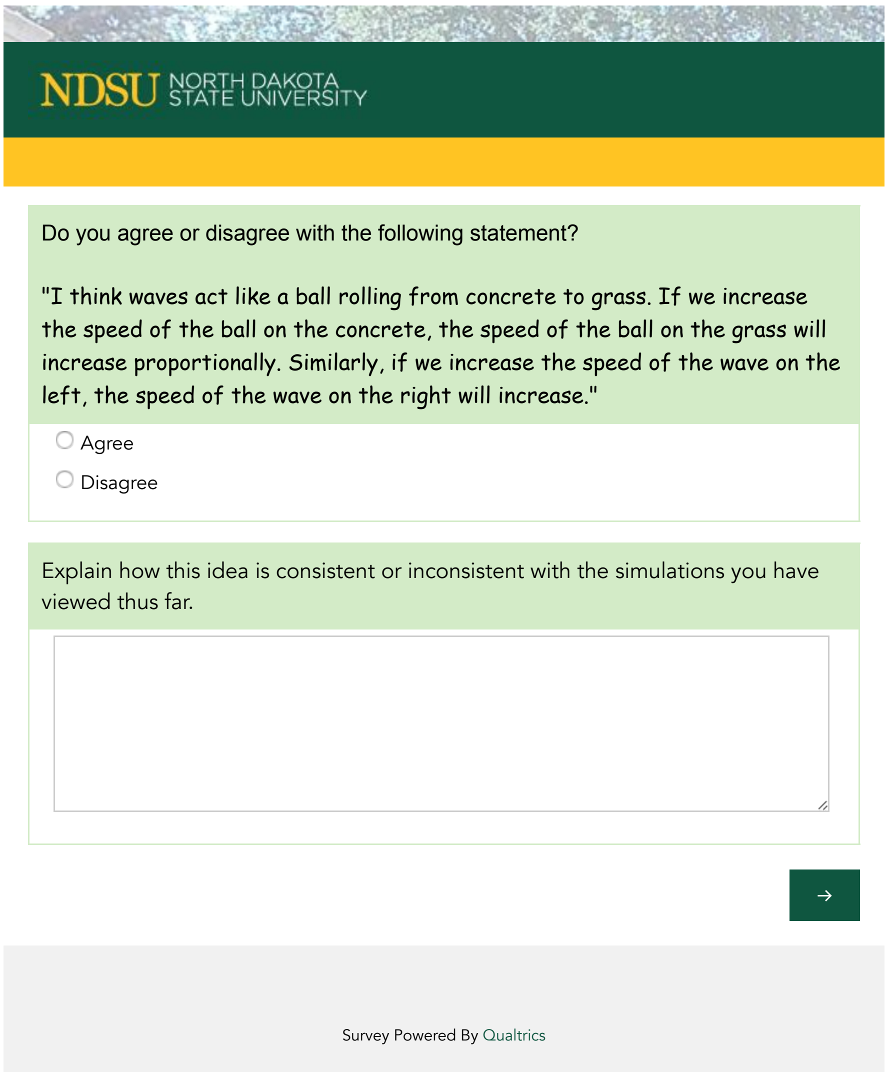



Figure A.2. Online mechanical waves lab (continued).


NDSU NORTH DAKOTA STATE UNIVERSITY

Do you agree or disagree with the following statement?

"I think waves act like a ball rolling from concrete to grass. If we increase the speed of the ball on the concrete, the speed of the ball on the grass will increase proportionally. Similarly, if we increase the speed of the wave on the left, the speed of the wave on the right will increase."

- Agree
- Disagree

Explain how this idea is consistent or inconsistent with the simulations you have viewed thus far.

→

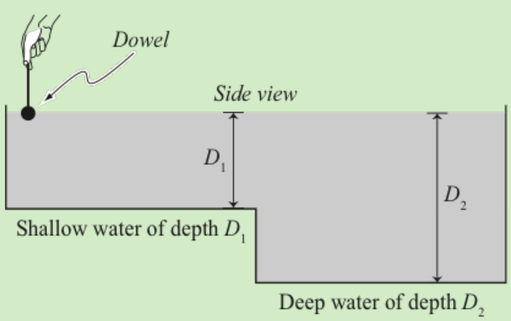
Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).

Part 4: Waves in a tank

Shown below is a side-view diagram of a large tank of water containing two regions of different depths. Water waves travel *more quickly* in deep water than in shallow water. A student generates a periodic wave (not shown) by a dowel tapping the water surface at a steady rate.

*L* and *R* refer to the left and right regions of the tank.




Consider the quantities below. Could changing any of the quantities *alone* result in the increased wavelength in the **left** region of the tank? Select all that apply.

- ☐ $v_L$
- ☐ $f_L$

Figure A.2. Online mechanical waves lab (continued).

Explain your reasoning.

Could changing any of the quantities *alone* result in the increased wavelength in the **right** region of the tank? Select all that apply.

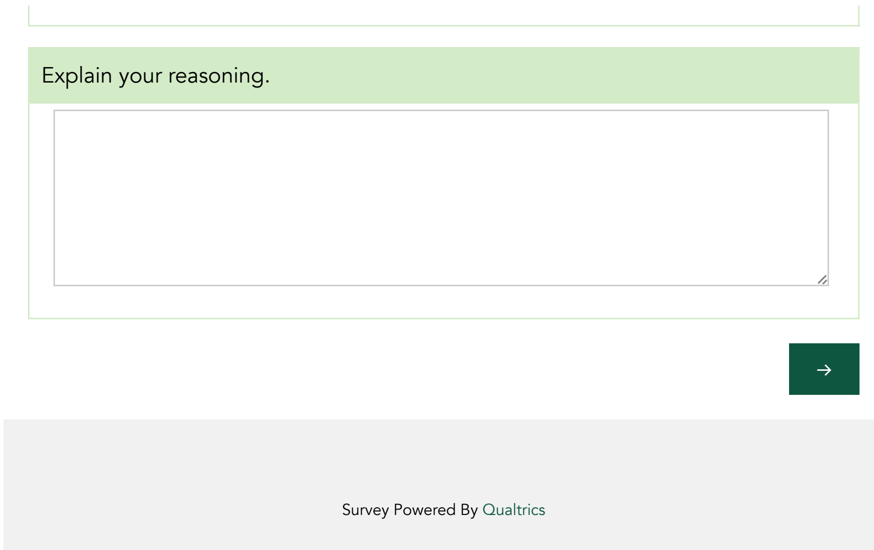
- ☐ $v_L$
- ☐ $f_L$
- ☐ $v_R$

Explain your reasoning.

Could changing any of the quantities *alone* result in the increased wavelength in the **right** region of the tank while the wavelength in the **left** region remains the same? Select all that apply.

- ☐ $v_L$
- ☐ $f_L$
- ☐ $v_R$

Figure A.2. Online mechanical waves lab (continued).


Explain your reasoning.

Survey Powered By Qualtrics

Figure A.2. Online mechanical waves lab (continued).