Novel Observations in Mixed Reality (NOMR): Designing a New Frontier of Physics to Practice Generating Scientific Models

Jared Phelps Canright

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Suzanne White Brahmia, Chair

Paula Heron

Elin A. Björling

Program Authorized to Offer Degree:

Physics

University of Washington

Abstract

Novel Observations in Mixed Reality (NOMR): Designing a New Frontier of Physics to Practice Generating Scientific Models

Jared Phelps Canright

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

The creation of new knowledge in the form of scientific models is a cornerstone of the process of science. In physics laboratory instruction, students are very often stuck in a confirmatory mindset; they have been conditioned to believe that the role of experimentation in the classroom and in the world is to verify known facts. This mindset is antithetical to the actual practice of science and detrimental to students' growth as scientific thinkers. Freeing students from such a mindset is a central challenge in the development and implementation of epistemologically authentic physics laboratory activities. This dissertation describes the exploration and implementation of a unique solution to this challenge: the complete removal of any and all traces of a "correct" answer through immersion in a different universe with its own laws of physics.

The Novel Observations in Mixed Reality (NOMR) project simulates real and fictitious laws of physics in a virtual reality universe to create learning opportunities where the only right answer is the one that students can make the best case for. NOMR's physics focuses around field-mediated particle interactions: familiar ones such as electrostatics and Newtonian gravitation, and novel ones created for students to explore from scratch. This dissertation reports on the development of the VR software, fictitious physical phenomena, classroom implementation practices, and instructional materials comprising NOMR labs. These labs are implemented across four large-enrollment courses on two WA campuses, reaching over 2,000 students each year. Assessment of NOMR's

effects on affective components of students' learning reveals that students become more expertlike in their epistemology and self-efficacy about experimental physics, and that they become more engaged with successive NOMR labs even beyond the expiration of the VR novelty effect. NOMR has been shown to be feasible and work as intended; it is primed for the next generation of developments and inquiry, building on the foundation presented here.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

NOMR: Novel Observations in Mixed Reality is the virtual reality physics laboratory platform developed, implemented, and evaluated over the course of the PhD project documented by this dissertation.

VLE: The Virtual Learning Environment refers to the VR application software built as part of this project. The software immerses the user in an interactive 3D virtual laboratory workspace containing familiar and novel physics and the tools to explore them.

VR: Virtual Reality is a fully visually immersive medium, where the user's view of the world around them is completely blocked and replaced with a computer-generated virtual environment.

AR: Augmented Reality refers to technologies that overlay computer-generated visuals onto a view of the real world.

MR/XR: Whether there is a line between Mixed Reality and eXtended Reality is a subject of some discourse. Both can safely be considered interchangeable umbrella terms for all forms of immersive technology, including AR and VR technologies.

ISLE: The Investigative Science Learning Environment is an approach to physics learning pioneered by Eugenia Etkina and colleagues at Rutgers University. At its core, ISLE aims to teach students to be physicists above teaching physics to students. The ISLE approach centers around the ISLE process of hypothesis-generating, hypothesis-testing, and hypothesis-application experiments, a simplified but authentic representation of the real-world process of science.

MODEL: is interchangeable with *Hypothesis* in context of ISLE and in this work. The term refers to a simplified conceptual representation of a real object or phenomenon. In physics contexts, this representation typically has descriptive or explanatory power, predictive power with limitations, and often includes physical properties represented with quantitative variables and related with mathematical equations [19].

ACKNOWLEDGMENTS

I have been supported, inspired, and kept sane by the combined efforts of countless friends, family, and colleagues over the last pandemic and a half. To fully acknowledge them all would require a full chapter unto itself. You know who you are, and you have my heartfelt thanks. It has been a delight and a privilege to learn, grow, and thrive with you all.

To have a graduate student come in with an idea for a brand-new project, disconnected conceptually and materially from any other project a new tenure-track faculty member is working on, and ask that faculty member to advise them on it, is a tall and unusual ask in physics education research. Thank you, Suzanne, for entertaining that ask and seeing it through to completion with me. I have become a better teacher, researcher, and collaborator through your counsel and your unwavering commitment to making every physics classroom a welcoming and enriching space.

I am deeply fortunate to have been part of the Physics Education Group (PEG) during my tenure at UW. Many details of NOMR's design, big and small, are directly attributable to conversations with and feedback from my colleagues in PEG. Every one of the works produced in conjunction with NOMR has benefited greatly from their insightful feedback. Thank you to Suzanne White Brahmia, Paula Heron, Peter Shaffer, and Donna Messina for maintaining this collaborative and supportive space. Thank you to my fellow (past and present) graduate students Anne Alesandrini, Jesse Ashworth, Dean Bretland, Cameron Flynn, John Goldak, Lisa Goodhew, Qirui Guo, Taylor GurrEithun, Ella Henry, Kristin Kellar, Aswin Rangkuti, Al Snow, Bert Xue, Sheh Lit Chang, and Charlotte Zimmerman for your boundless advice and encouragement.

I cannot express in words my appreciation for the tireless passion and effort David Aplin and Eddie Mendoza have poured into NOMR's continued development and implementation. From helping make the force meters (finally) behave, to shouldering headset management, to writing

STF grants, to building the NOMR Student Developer Group from scratch, you have gone above and beyond at every turn to make this wild little idea of mine into a community and a core part of UW PHYS lab curriculum that will endure past my tenure at UW. NOMR is as much yours as it is mine, and I could not be more proud to share it with you.

NOMR has benefited from the contributions of many undergraduate students over the past several years, several of whom continue to work on it to date as part of the Student Developer Group (SDG). I thank every student who has contributed to the NOMR labs:

PHYS 499 students: Adin Pierce, Hayden Potter, Darcy Chen, Michael Raquel, Tyler Flom, James Nelson, Adam Tazi, Boris Chkodrov, and Igor Banin.

NOMR Student Developer Group: Leonardo Chen, John Delker, Will Erhard, Shiqi He, Bryce Larson, Tony Le, Charlie Norgaard, Nathan Parece, Roberto Raftery, Jonas Rivera, Alnis Smidchens, Gary Tong, Felicia "Fish" Tsai, Sanjali Vuriti, Ian Wang, and Maggie Zheng.

Special thanks go to Prof. Kazumi Tolich for her great intellectual contributions to the NOMR lab manuals, TA guides, and TA trainings. It is thanks to you and the physics lab instructional team that the NOMR lab activities have grown from a fun research idea to a mature series of activities that can be deployed each quarter and across institutions.

Many aspects of NOMR's design, use, and evaluation were guided by thoughtful discussion and advice from members of the broader PER community. Special thanks to Drew Rosen, Eugenia Etkina, Andrew Boudreaux, and many more for your time and input.

I found being invited for and giving colloquium talks to be the most professionally validating experiences in my graduate school career. Actions speak loudly, and I cannot imagine a louder way of saying, "I see you as a fellow scientist, and you are doing work that I and my colleagues find interesting." Thank you to Joel Franklin at Reed College, Lisa Young at New Mexico Tech, and Andrew Boudreaux at Western Washington University for giving me these opportunities.

I attended my first Out in STEM (oSTEM) conference in Denver in 2016. That life-changing experience led me to volunteer with the organization, which introduced me to a community and

friends who I will forever hold dear. oSTEM has become central to my identity as a queer person, as an education researcher, and as a member of the STEM community. I thank oSTEM and all my fellow volunteers for keeping me going with the annual infusion of queer joy at each conference, and for the friends, breadth of perspective, and opportunities for professional growth it has brought me.

This wouldn't be a STEM dissertation without one or more key high school figures who inspired me to pursue this path. Thank you to Mr. Hansen, Ms. Brown, and Ms. McCarthy at Hutchison High for sparking and fostering my love of the sciences. Thank you to Stephanie Rudig at Weller Elementary for getting me into MathCounts and LEGO robotics even before I knew the word "physics."

2020 and the ensuing lockdown was a rough time, made much less so by having people to weather the storm with. Audrey, Kato, Josh, Anthony, Dan, and Anna, thank you for making that time in our lives memorable in a good way.

Last but certainly not least, I thank my family for creating the foundation that allowed me to reach this point: Mom, Dad, and Jacob. I drew all of my best traits from you, and I look back on my upbringing as a homeschooled kid in Alaska with fondness and pride. Josh, my love, I cannot imagine a better person to spend a PhD, a pandemic, and the rest of my life with. I do not know that I could have finished this PhD without you and Dahli to come home to every day.

DEDICATION

To Dahli, my dear sweet girl, the cantankerous wench who rescued me as much as I did her. Even though you ate the last of the NOMR cookies Charlotte made for my defense. Rude.

Chapter 1

INTRODUCTION

A brief foreword: This dissertation describes the development and evaluation of an interactive virtual reality learning environment. Capturing the in-headset experience is impossible with static text and images. At the following permanent URL, one can find links to video demonstrations of this project, its current web presence, contact information, and links to any public version which may later be published, which will be updated as the project continues to evolve. https://purl.org/nomr/dissertation

Teaching students to think like physicists is a key goal of physics education. Science is built upon an iterative process of creation and validation of knowledge from empirical data; understanding and learning to practice this process is a crucial step in the training of a budding scientist, and an important takeaway from science class for members of an informed public.

Yet, introductory physics courses at the college level generally do not help students become more expert-like in their understanding of the role of experimentation in science, nor more confident in their ability to practice experimental science [48, 56]. Students tend to adhere to confirmatory mindsets in physics labs, even in the face of direct instruction to the contrary, as they are overwhelmingly conditioned to understand the role of experimentation in class as a way of confirming known facts [48, 26, 49].

The Novel Observations in Mixed Reality (NOMR) project approaches this challenge by creating learning opportunities where the only correct answer is the one that students can make the best case for. By using virtual reality (VR) to immerse students in a new universe with its own unfamiliar laws of physics, a learning environment is created where there is no access to any "correct" answer whatsoever. In absence of a canonical expected result, students have nothing to confirm or be judged against, allowing them to engage in the authentic practice of science as a means of

creating knowledge from empirical observations.

At a glance, NOMR is a virtual reality physics lab platform that simulates real and fictitious physical phenomena in the vein of electromagnetism and gravitation; all of NOMR's physics can be expressed as force laws governing interactions between point particles. It is used in 100-level and 200-level physics lab curricula at the University of Washington (UW) and Washington State University (WSU) campuses, where it comprises 3-6 weeks' worth of lab activities in a course.

This dissertation describes the development, implementation, and assessment of NOMR's VR software, instructional materials, and classroom implementation practices.

To begin, Chapter 2 describes the what and why of NOMR: Its purpose, its pedagogical underpinnings, the constraints that guided its design, and the instructional content that comprises it.

Chapter 3 describes the who, when, and how of NOMR. It is a six-year chronological account of the development of NOMR's precursors, the NOMR virtual learning environment (VLE) software, the instructional material and classroom implementation practices enabling its use, and the community supporting it. NOMR's precursors in the realms of special relativity and Gauss' law provide the foundation for the first prototype and pilot test of NOMR. Classroom safety procedures, VR locomotion and interaction design patterns, and technical implementation of the NOMR physics engine are covered in detail. Lessons learned from each implementation of NOMR give way to a stable routine of classroom implementation independent of my involvement, and my focus turns to education research on NOMR's effects.

Studies of NOMR's affective effects are described in Chapter 4, itself a standalone paper soon to be published in Physical Review: Physics Education Research. NOMR's effects are assessed along axes of students' epistemology about experimental physics, self-efficacy in experimental physics contexts, and engagement quality with each lab exercise. Gains are observed along all three axes, though the studies fall short of characterizing the robust, energetic creativity we see students engage in with NOMR labs.

This work is deeply exploratory and interdisciplinary, spanning physics education research (PER), physics curriculum design, VR development, human-centered design, educational technol-

ogy, and adjacent fields. As such, this dissertation contains elements of all of these fields . I break down which elements may be of interest to a few archetypical readers below:

A **physics educator** interested in learning more about NOMR would be well-served by Chapter 2. How NOMR is used in the classroom is covered in Section 2.6. The NOMR lab activities used at UW are described in detail in section 2.7; the complete lab handouts are in Appendix B. Detailed examples of the student experience of NOMR are described in Section 2.5, complemented by example lab reports in Appendix A and an earlier proceedings paper in Appendix C reporting on case studies of the pilot implementation of NOMR.

A **curriculum developer exploring fictitious physics** will want to read Section 2.4 on the principles behind the development of NOMR's fictitious physics. For the process by which a couple of examples of fictitious physics were developed, see the latter half of Section 2.5. I encourage anyone interested in developing instructional activities with fictitious physics to contact me; there are many lessons learned that did not make it into this dissertation.

A developer of VR/AR education applications interested in the lessons learned from NOMR's development will find them in Chapter 3. I also recommend that they peruse the reflections on the role of VR in physics education in Sections 4.6.4 and 4.7.

A **physics education researcher**[1] interested in the evidence backing NOMR's fictitious-physics approach to inquiry labs should jump to the evaluation of NOMR in Chapter 4, which is soon to be published as a standalone paper. Motivation of the fictitious-physics approach can be found in Sections 2.2 and 2.3. Proceedings papers produced in conjunction with this dissertation are included in Appendix C.

[1] A physics education researcher might note that this dissertation is more akin to an experimental physics dissertation than to a PER dissertation, in that it describes creating something for 3-5 years, followed by 1-2 years of doing science with that thing. That assessment would be spot on.

Chapter 2

DESIGN

This chapter describes the what and why of NOMR: Its purpose, its pedagogical underpinnings, the constraints that guided its design, and the instructional content that comprises it.

Sections 2.1, 2.2, and 2.3 describe relevant background literature, constraints on and goals of the initial development of NOMR, and the core design principles underlying NOMR, respectively. Section 2.4 goes into detail on the process of designing novel laws of physics for NOMR. Section 2.5 describes two sets of novel physics, or "scenarios." Each scenario is introduced through an example of students' exploration thereof, and followed by a description of the development of the programmed rules underlying each scenario. Section 2.6 discusses the role of the instructor in NOMR-powered lab activities. Section 2.7 gives a detailed example of how NOMR is currently used in PHYS 122 (Introductory Electromagnetism) and PHYS 231 (Introduction to Experimental Physics) lab curricula at UW. The lab handouts associated with this example are included in Appendix B.

2.1 Background

2.1.1 ISLE

The Investigative Science Learning Environment (ISLE) is an approach to physics learning pioneered by Eugenia Etkina and colleagues at Rutgers University [16, 5, 14, 15, 17, 6]. In brief, rather than teaching physics to students, ISLE aims to teach students to be physicists. Students' understanding of and comfort with the process of science as a means of creating and validating new knowledge takes priority over physics content coverage.

In a full implementation of ISLE across lecture, lab, and recitation sections, students cover new physics content through the ISLE cycle. Beginning with hands-on observational experiments,

students generate explanations and winnow them down into models, which are subsequently tested and refined. This process of knowledge creation and validation is the means by which students learn physics.

ISLE Process

Every ISLE activity follows the ISLE process, a simplified but authentic representation of the process of science, as in Figure 2.1.

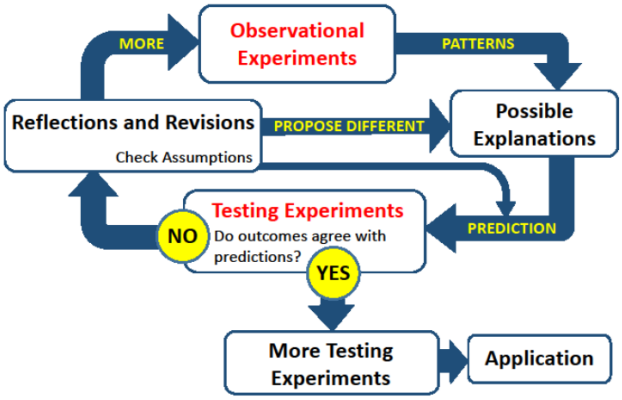


Figure 2.1: The ISLE process is a simplified representation of the real-world practice of science, iteratively generating, testing, rejecting, and refining models to empirically create and validate knowledge.

The ISLE process is one of iterative creation, evaluation, and refinement of scientific knowledge through experimentation. Definitions of key terms follow:

Model is interchangeable with "hypothesis," and refers to an explanation for some physical behavior of some system. A model must be testable, falsifiable, and have predictive power.

Observational experiments, also called model-generating experiments or empirical modeling experiments, are used to create models. They are a place for open-ended exploration and rapid iteration over many ideas–often the wilder the better.

Testing experiments are used to test models. A model is used to create a prediction about how a

physical system will behave under some conditions; that prediction is then tested by observing the system in question under said conditions. Unlike an observational experiment, the outcome of a testing experiment is binary: Either the experiment rejects the model, or it fails to reject the model.

Application experiments use a tested model to solve a real-world problem or measure a quantity.

ISLE is driven by two core intentionalities which can also be considered to be pillars of NOMR's design [5]:

1. **Students should learn physics by engaging in activities that mimic the authentic knowledge-generating activities of practicing physicists.**
2. **All our choices of theoretical perspectives and our curricular design decisions should enhance or nurture human well being rather than harm it.**

ISLE Example

Johannes Kepler's laws of orbital motion are an illustrative historical example of the ISLE process. His work took place before Newton's time, absent a theoretical motivation (gravity) for the motion of celestial bodies. Using a collection of Tycho Brahe's empirical data of the motion of Mars, Kepler determined that planetary orbits are not circles, but ellipses. Without the notion of gravitational attraction to guide his work, he nonetheless derived mathematical expressions relating the period of an object's orbit to the geometry of its orbit. Through the lens of the ISLE process, we would consider Kepler's work to be an observational experiment, and his three laws a possible explanation, or model, of the phenomenon of planetary motion.

Kepler's laws were derived from the motions of Mars, and underwent model-testing experiments through examination of the orbits of other planets. In each case, Kepler's laws held. In 1687, roughly 70 years after Kepler's final publication, Isaac Newton published his seminal text, *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), containing Newton's three laws of motion. Newton's law of gravitation $\left(F = \frac{m_1 m_2 G}{r^2}\right)$ was presented as a case study of the laws of motion; from his theory of gravity, Kepler's laws could be derived and explained. In this way, Kepler's model was integrated into a more expansive model,

opening up the doors for more testing and application experiments. The Cavendish experiment was the first to empirically determine a value for the gravitational constant G in Newton's law of gravity, which we can consider an application experiment using Kepler and Newton's combined model [43, 38, 32].

For a thorough example of an application of the ISLE process and approach in the classroom, I urge the reader to peruse Eugenia Etkina's Millikan award lecture paper [15]; it is open-access, easy to read, and does a much better job of illustrating the practices and mindset underlying ISLE than any paraphrased version I could produce here.

2.1.2 Inquiry Labs

Need for epistemologically authentic inquiry labs

Creating physics lab curricula that teach students how the scientific process is used to create and validate new knowledge in the same way professional scientists do is an open problem in physics education. A collection of studies point out that the content-focused "cookbook" labs prevalent in physics education over the past decades neither improve students' understanding of the conceptual physics knowledge in question [24] nor aid in their growth as budding scientists [56, 48]. Trumper's 2003 paper [53] puts the issue succinctly:

> " Laboratories of this type are better omitted from courses because they discourage students, offer no new information about nature, and even present a wrong view of the process of science. "

In particular, there is a need for opportunities for students to practice the creation of scientific models [31, 19]. While there are exceptions, physics lab curricula in large-enrollment courses generally do not reflect the iterative nature of scientific inquiry, leaving students with a poor understanding of the complex interplay between experiment and theory involved in the creation of new scientific models [27, 48].

Student adherence to confirmatory mindsets

Students in undergraduate physics labs are overwhelmingly conditioned to believe that experimentation in class is meant only to reinforce physics content or confirm a "correct" answer [50, 52, 49, 25, 27].

In-class experiments are most often thought of as exercises in confirming what is already known, and experiments in professional science are seen only as a means to test theories, not create or develop them. Students see experiments as interactive demonstrations, not as a way of learning in class, and not as a way of knowing in professional science.

In the presence of known answers, students tend to hold those in highest regard, seeking to confirm known answers above anything and everything else, even in the face of contradictory data [52]. It is common practice for students to go so far as to engage in questionable research practices such as inflating error bars and discarding disagreeable data in order to achieve agreement with theory.

Great creativity is required to design lab activities in such a way that students don't see their findings as something that has been discovered before and can be Googled. With Google's ubiquity, there are precious few opportunities for students to think creatively about an unfamiliar system and create an empirical model of it through experimentation. Instructors and students commonly recognize that any physics that students might be asked to discover in a lab has been discovered before and is easily found on the Internet.

Approaches to inquiry labs

Bartlett and Dunnett [4] describe the "secret objectives" approach to inquiry labs, focusing on inconsistencies between experimental reality and the information students are given about an experiment. Students are explicitly challenged to identify and investigate secret objectives in each laboratory activity, engaging in genuine inquiry activities in doing so. One example of a secret objective is the damping of a simple pendulum. Each activity may have more than one secret objective, and not all such objectives need be known by the teaching staff. Rather, it is important that

the instructor's knowledge of secret objectives in an activity remain short of comprehensive:

> " It is important to acknowledge that secret objectives are personal and that the teaching staff only know some of them to discourage students from believing that there is a correct answer - or that there is a particular point that the demonstrator team are expecting them to find and investigate. "

Allie, Buffler, Kaunda, and Inglis [2] describe laboratory activities in which students' laboratory tasks are presented as accounts of fictitious experiments in the campus newspaper. Presenting the lab prompts in this way helps to negate the authoritative nature of instructions and promote students' agency. In these accounts, an experiment that produced a result discrepant with accepted theory (e.g., a measurement of the gravitational constant $g \sim 16$ m/s^2) is described. Students are tasked with determining the source of the discrepancy and conducting and reporting on a similar experiment of their own devising.

As part of a customer discovery effort when I was thinking about commercializing NOMR, I interviewed roughly a dozen experienced high school physics teachers about their approach to inquiry-centric learning. Anecdotally,[1] the message I gleaned is that experienced teachers develop a classroom culture that centers and fosters students' own agency, and this culture is sufficient to keep students out of a confirmatory mindset. Every teacher has their own way of doing so, developed over years of experience, and a new teacher copying their course materials would not necessarily be able to achieve the same effect.

2.2 Constraints and Synthesis

This is the awkward middle child section, combining contextual constraints with my own motivations, musings, and ideas to make connections between the literature (Section 2.1) and the design principles core to NOMR (Section 2.3. In other words, this is the reference-free counterpart to the previous section, in which all of the non-literature ideas, motivations, and realities underlying NOMR's design principles are discussed.

---

[1] Read: I do not have a reference for this paragraph, so take it with a grain of salt.

2.2.1 Context: Transformed lab sequence

This subsection describes the transformed introductory physics lab sequence of which NOMR is a part, and from which it inherits a great deal of pedagogical structure and purpose.

The transformed physics lab sequence (TLS) at UW is designed for use as the three-quarter introductory physics lab curriculum used by the 100-level calculus-based and algebra-based course sequences. The same lab curriculum is used by both courses, though the algebra- and calculus-based students have separate lab sections from each other and their grading structures differ slightly.

Due both to the timing of its development and the existence of assets from a previous VR lab setup, NOMR ended up being developed for use in the second-quarter lab curriculum, falling under the umbrella of electromagnetism.

The reason we say the transformed labs (and thus NOMR) *align* with ISLE is that it is not meant to be used as a part of or alternative to a full ISLE implementation. Rather, the transformed labs use ISLE as a pedagogical framework upon which to build an epistemologically authentic learning environment.

Every lab in the TLS can be categorized as a model-generating, model-testing, or model-application experiment, and often a model-generating experiment one week feeds into a model-testing experiment in the next.

Create opportunities for students to create new knowledge

Writing model-testing experiments is a sight easier than writing model-generating experiments. The latter require a great deal more TA buy-in and training, as most incoming graduate TAs come from a background of confirmation labs. The confirmatory mindsets of both students and TAs must be fought every step of the way. Writing an observational experiment that is both tractable and open-ended is an exercise in frustration, as students latch on to anything they can find on the Internet that they think their instructors want them to end up with. Even if students refrain from looking up their system during lab, it is very likely that they will do so as they are writing their report.

It is in the midst of experiencing this specific pedagogical frustration that the ideas behind NOMR came about. While in its formative phase, NOMR's development was driven by a need to create opportunities for students to create new knowledge, in the form of model-generating experiment activities appropriate for an introductory electromagnetism context.

2.2.2 Motivation: Effectively use VR in the physics classroom

It would be accurate (if reductive) to say that my desire to find a niche in physics education where VR can be uniquely useful is what drove the creation of NOMR. Previous large-scale controlled studies comparing the effectiveness of interactive VR, interactive 2D simulation, and 2D video treatments of the same physics content [51, 40, 33] have consistently found that the VR treatment group did not perform significantly better than the rest. My own experience with a VR lab teaching flux and Gauss' law (Section 3.2) produced the same finding.

I interpreted these findings to mean that even for inherently 3D phenomena, having students interact with visualizations of physics content in VR does not significantly improve their learning on its own. I hypothesize that if a VR experience can be readily converted into a 2D simulation or video format with little to no loss of ease of use or presentable content, then the affordances of VR were never in use by that experience.

After my own null result, I still held a strong belief that there had to be a niche where VR could be useful in the physics classroom. After internalizing these findings, I decided that the feasibility of a controlled comparison of VR to non-VR interventions could not be a factor in the design of future prototypes. At the cost of making the VR-powered activities more difficult to assess, dropping this requirement opened up a great deal more room to be creative and make something that we could not do without VR.

2.2.3 On confirmatory mindsets and "correct" answers

Arguably one of the greatest difficulties in inquiry-focused physics laboratory instruction is the presence of a "right" answer. Knowing the canonical answer, how to find the answer, or even that

such an answer exists somewhere lets students latch onto the idea that their role is to confirm some external authority's pre-ordained "right" answer. By and large, students are strongly conditioned to believe that such an answer exists and that the goal of experiments in the classroom is to confirm that answer. Even in situations where their grade is very clearly independent from their experimental results, students will fudge their numbers to agree with the textbook. This confirmatory mindset is antithetical to an expert-like understanding of experimentation in physics and a barrier to laboratory instruction that centers students' development as practitioners of science.

Shifting students out of this confirmatory mindset is not an insurmountable barrier, but it is a complex one that generally requires significant pedagogical expertise and experience on the part of the instructor to overcome. Section 2.1.2 describes a couple of anti-confirmation approaches. It is difficult to consistently cultivate a classroom environment where students feel safe putting their confirmatory mindset aside. This difficulty is exacerbated in large-enrollment introductory course contexts where students' laboratory instructors are relatively inexperienced graduate teaching assistants. Figuring out how to write curriculum and train teaching assistants in such a way as to cultivate such an environment is an active area of physics education research on laboratory instruction. The remainder of this dissertation is my contribution to this effort, centered around the idea: What if there were no "correct" answers at all?

2.3 NOMR Design Principles

2.3.1 Create a space where there is no correct answer

NOMR attacks the problem of epistemologically authentic physics laboratory instruction from a unique angle: **Completely remove any and all trace of a canonical correct answer.** This is admittedly an extreme means of disabusing students of the notion that there is a "right" answer, but if this does not help, I truly do not know what will. Keeping in mind that even knowing that a "right" answer exists can keep students in a confirmatory mindset, taking this approach means that documentation of the phenomena students are exploring cannot be present on the Internet, and neither the students nor their instructors should have access to any kind of "answer key." **In both**

the students' and instructors' eyes, the only right answer should be the one that students can make the best case for.[2]

To completely remove all trace of a canonical correct answer, allow for a range of viable experimental outcomes, and do so at a level approachable to introductory physics students leaves us with precious few, if any, real physical phenomena that can be explored with hands-on lab equipment. Thus, we turn to AR and VR to create our own universe, with our own novel physics, custom-built from the ground up for students to discover. In the words of the wise and mighty Thanos [44]: "Reality can be whatever I want."

For all instructional intents and purposes, the programmed rules governing the novel physics do not exist. They have no place in the instructional practices around NOMR, much less the interpretation or evaluation of students' findings. Even if those rules come to light, it should be emphasized that they are not "correct" answers, as they do not fully explain the behavior of NOMR's physics: the real-time game physics engine underlying the NOMR application generates many behaviors that are not explained by these rules. Students' experimental findings are the highest authority on NOMR's physics.[3]

By corollary, there should be room for multiple valid experimental outcomes from examination of a given phenomenon. If there is "no right answer" but every group in the class except one arrives at identical conclusions, that one group is more likely to reanalyze their data until it fits the class's than to interpret the discrepancy as the opportunity for further inquiry and scientific discourse that it is. The more defensible variance in students' findings, the better.

In short, NOMR aims to create a frontier of physics that is tractable for introductory students to explore. Creating a space where there is no right answer serves to emulate the experience of real-world research, thus aligning with ISLE's core principle: Teach students to be physicists.

---

[2]Incidentally, this also means that ChatGPT and friends cannot possibly know anything about NOMR's physics. For better or for worse, AI resilience is now probably one of NOMR's strongest selling points.

[3]To support this notion, students are encouraged to name the particles they discover–a strongly expert-like behavior.

2.3.2 Fit into introductory electromagnetism lab curriculum

Without going into great detail (see the process chapter for that detail), NOMR was originally created to use in the transformed introductory electricity and magnetism labs at UW. Thus, a major guiding force in its design was the need to draw from similar physics concepts as the labs which might otherwise take its place. As the transformed labs are heavily aligned with ISLE, specific content coverage takes a back seat to teaching students to be physicists, but NOMR still had to fall within the umbrella of electricity and magnetism. Thus, NOMR focuses on particle interactions, using Coulomb's law as an approachable segue into more exotic physics. In doing so, NOMR fits nicely alongside a series of circuits labs to form a quarter-long electricity and magnetism lab curriculum, as discussed in Section 2.7.

A potential pitfall in VR science lab platforms is pedagogical inflexibility: Specific activities are programmed into the VLE, leaving little room to adjust the learning goals and structure of the students' lab activity in response to changes in the lab curriculum writ large. NOMR was developed before the transformed labs had reached maturity. The NOMR VLE had to be flexible then to continue to have a place in the ever-evolving transformed lab series; today, the same flexibility is necessary to bring it to other campuses and instructional contexts. Thus, it is designed to be a sandbox largely bare of prompts or instructions, not unlike PhET simulations. This way, a breadth of lab activities can be designed to make use of the same VLE.

2.3.3 Maintain affordances for teaching uncertainty, etc. inherent to real-world research

Development of experimental data collection and analysis skills, focusing on first-principles development of techniques for handling experimental uncertainty, is a core element of the transformed introductory lab sequence at UW. Thus, NOMR must support the growth of students' knowledge of quantifying experimental uncertainty and its role in the interpretation of experimental results.

Specifically, this means that there can be no "perfect" data as one can often produce from a simulator. Each measurement made in NOMR must have experimental uncertainty associated with it. In a PhET, for example, one usually obtains identical results with identical inputs, and

there is no barrier to setting up those identical inputs repeatedly. In a real-world measurement, say, the measurement of the length of a table, taking the same measure multiple times will yield different results. Different experimenters (or the same experimenter at different points in time) might perform the measurement slightly differently in any number of little ways. Variations in how taut a measuring tape is pulled, which point on the table is measured, how the experimenter handles the slightly rounded edge of the table, etc. are all examples of the innumerable little sources of uncertainty inherent to real-world measurements.

Inserting those sources of uncertainty into a 2D simulator controlled by mouse and keyboard would likely be perceived as "janky" or "clunky," and generally feel like an unwelcome and unnatural addition. The closest one could get to a "natural" source of uncertainty in a 2D simulator would be in placement of a point via mouse or touch input. Even then, it is common to expect snapping behavior in mouse- and touch-based user interfaces to aid with exactly that. Were a simulator configured to produce uncertainty in this way, it would be seen more as a frustrating limitation of the simulator than as an inherent part of experimentation.

Thus, measurement instrumentation in NOMR must be set up to mimic natural sources of uncertainty in real-world experiments. The use of physical controllers tracked in 3D space or hand-tracking provides a straightforward means of doing so, as neither has the perfect precision achievable with mouse and keyboard, and that is something we're used to. Picking up and putting down an object in 3D space naturally has a bit of error to it, even if one is actively trying to put it back exactly where it was.

The need for a range of potential findings aligns well with the need to teach students how to quantify and interpret uncertainty in their experimental data and analysis. It is not uncommon for multiple functions to fit students' data within their error bars.

2.4 Design of novel physics

The physics content space covered by NOMR is dynamic field-mediated interactions between point particles. Particles are modeled as hard spheres in the VLE. NOMR's novel physics are created by defining force laws governing interactions between point particles. This section dives into the

process of developing new physics and considerations made while doing so.

2.4.1 Agreement with known physics

Making up new laws of physics for students to explore was absolutely one of the most fun parts of this project. It poses an interesting challenge, one I imagine is quite familiar to game designers: Designing a system of which one has full knowledge, for users to explore with no prior knowledge thereof. It is very near impossible for me to experience the NOMR phenomena in the same way students do, because I was the one to write down the physical laws dictating their behavior, and I am not terribly good at selectively forgetting things. In creating a NOMR scenario, one is creating the emergent behavior of the phenomena and the experience of observing that behavior with fresh eyes, more than one is creating the fundamental rules driving said behavior. I found traction on this problem by establishing firm boundaries on the physics design and focusing on creating complex behavior from simple rules.

In designing the novel physics used in NOMR, I came to recognize three rules that do a neat job of defining the boundary between what is consistent with known laws of physics, and what runs the risk of conflicting with students' existing understanding:

(1) Thou shalt not violate Newton's laws.

(2) Thou shalt not violate conservation of momentum or energy.

(3) Thou shalt not imply the existence of a universal reference frame.

Rule (1) came about immediately. We spend so much time getting students to understand the concept of third-law pairs that putting them in situations where Newton's laws do not seem to apply would be baffling at best and educational malpractice at worst. Newton's laws are conceptual bedrock that introductory students should be able to build their physics knowledge upon, and NOMR should do nothing to jeopardize that foundation. NOMR's physics are designed specifically for students to develop mathematical models thereof through experimentation. If they cannot assume Newton's laws to hold, this task becomes intractable in a hurry.

This rule constrains the design space by requiring that every pair of interacting objects be a

third-law pair.[4] Newton's second law further stipulates that we cannot horse around with $F = ma$.

Similar to Rule (1), violating Rule (2) would risk students' confidence in the universal applicability of energy and momentum conservation. Epistemological hazard aside, letting particle interactions create momentum or energy out of nowhere is a recipe for very strange and out-of-control behavior.[5] It is important to note that this rule applies specifically to particles and their interactions; the user and their tools are by necessity allowed to violate energy and momentum conservation.

For example, by anchoring a particle at a point in space in the VLE, one is robbing the system of the momentum and kinetic energy that the particle would be accruing in its interaction with other particles. In our universe, this would be akin to anchoring one end of a spring to a heavy object. Assuming the spring is sufficiently lighter than the object, momentum and energy are conserved through interactions with the earth. However, if the spring's mass were increased, it would eventually pull the anchor along with it, making those conservation laws more apparent. The same is not true in NOMR; the user and their tools are treated as having infinite mass, and no amount of force will move the anchors or the user's workspace. This is an example of where the distinction between *true* and *apparent* physical accuracy comes into play, as discussed in Section 2.4.3.

Rule (3) is a bit more subtle and emerged less from concerns over harming students' conceptual understanding than from my experience that any scenario which violated this rule felt cheap:

- A force that operates along only a subset of the cardinal directions, and ignores others.
- A force that is stronger along some cardinal directions than others.
- A force that depends on the velocity of the particle in the lab frame.

These ideas did not pan out to be satisfying. Coupled with the epistemological hazard, I decided this was a useful and necessary constraint on NOMR's physics.

---

[4] In context of non-Galilean relativity, this need not necessarily be true. However, implementing a full special-relativistic particle interactions engine was well out of scope, as interesting as the result might be.

[5] Real-time physics engines play fast and loose with conservation laws to begin with; exacerbating that problem was not a successful basis for scenario design.

2.4.2 What is a scenario?

A *scenario* is the basic unit of content in NOMR. A set of distinct particle types, the properties intrinsic to those particles, and the force law(s) giving those properties meaning comprises a scenario. For example, the electric interactions scenario consists of positively charged particles and negatively charged particles of equal masses and charge magnitudes which interact according to Coulomb's law. One might create another scenario by, for example, changing the r-dependence of Coulomb's law to something other than $\frac{1}{r^2}$, or offering other particle types with different charge and mass values.

The physical consistency rules above constrain the scenario design space to that which is acceptable within the laws of physics, but leave a great deal of room for phenomena that are needlessly obtuse, difficult to make heads or tails of, or just unfair. A couple of additional guidelines help to smooth over those rough edges:

1. Every interaction must be explained completely by the presence and behavior of other visible objects in the space.
    - Explicit time-dependence is off the table.
    - No invisible particles.
    - Turn off gravity, and create no invisible, externally-imposed global fields.
2. Less is more. Aim to create simple rules that give rise to interesting emergent behavior.
    - A simple force law does not by any means necessarily equate to a simple investigation thereof.

In designing a scenario, it is important to have some low-hanging fruit, as it were. That is, students should be able to identify *some* aspect of the scenario's behavior with relative ease. For example, students asked to create a model for the interactions between bar magnets are likely to come to the conclusion that there are two distinct poles; that is the low-hanging fruit. Their investigation can go on to look at more complex ideas such as the strength of the magnets' interaction as a function of separation, but even if they get no farther, they can talk about how many distinct poles there are and how they came to that conclusion.

At the same time, each scenario should have depth. In other words, there should be multiple distinct questions one can ask about its behavior which will yield interesting answers. If a group keeps asking questions and digging deeper, there should be more to find. In this we diverge a bit from the goal of emulating real-world research, where null results and boring answers to interesting questions are par for the course. Scenarios should specifically be designed to reward creative experimentation and exploration.

2.4.3 True and apparent physical accuracy

A helpful distinction to make in considering the design of NOMR's physics is that between *true* physical accuracy and *apparent* physical accuracy. No matter what interactive 3D engine NOMR is implemented with, the real-time physics simulation underlying it will be in essence a glorified numerical Euler simulation. These have error. Lots of it. High-precision physics engines used in research and engineering minimize this error in order to match the real world as closely as possible, at the cost of being unable to run in real time. Real-time physics engines designed for use in games (e.g., PhysX and Havok) have no obligation to *precisely match* the real world; their goal is to produce results in real time that *appear to be consistent with* the real world in a processing- and memory-efficient manner. Real-time engines can cut all kinds of non-physical corners in the background to save performance, so long as the results they produce are plausibly physical.

All this to say: There is error and non-physical behavior inherent to any interactive 3D physics simulation, and choosing to lean into that is the only road that does not lead to madness. It is inevitable that, through deliberate choices and otherwise, there will be violations of the laws of physics apparent to anyone who goes looking for them. What ultimately matters is the measurements and observations that the end user will see. All numerical measurements should yield results consistent with the laws of physics. For example, measuring the force two particles exert on each other should not give results inconsistent with Newton's 3rd law. Any kind of wild abuse and neglect of the laws of physics in the backend goes, so long as what the end user sees is apparently physically accurate.

2.4.4 Fictitious physics vs. esoteric physics

Multiple colleagues have noted that one could use physics that exists (or is theorized to exist) in the real world but is esoteric enough that introductory students would not recognize the resulting behavior. For example, particles interacting according to the Yukawa potential, or particles with decoupled gravitational and inertial mass and different ratios thereof. I think this is a great idea. However, it does raise the question: Should students be told at the end that they have been discovering "real" physics, and would creating a reference to canonical correct answers by doing so run the risk of defeating the purpose of NOMR? Likewise, one would need to be careful about NOMR's use in upper-division courses, as more advanced students might recognize, e.g., the Yukawa potential. Even so, I do not think these concerns are great enough to preclude inclusion of esoteric physics in NOMR. The primary reason it has not happened is avoidance of scope creep.

Prof. Masha Baryakhtar in the UW Physics Department and I have had a couple of conversations about incorporating the development of novel physics for NOMR as a class project for the graduate quantum field theory (QFT) course she teaches. Exploring physics beyond the standard model (BSM) is a common theme in QFT class projects; she proposed that students could begin with the QFT picture of a BSM theory, boil that down into the theoretical particles and interactions implied by the theory, and have an interpretation of those particles implemented in NOMR such that one can physically explore the implications of the BSM theory. I think this is a very promising avenue for future development.

2.4.5 Why VR?

The honest answer to this question is that I came to graduate school wanting to do something at the intersection of immersive technology and physics, and stubbornly pursued that notion until something panned out.

However, giving a real answer to this question is a useful lens through which to consider NOMR's design. Any educational application incurring the expense and implementation overhead of VR should be able to answer this question; otherwise it is likely not making use of the

unique affordances of VR.

This section is written more or less completely from my experience as a teacher, VR developer, and VR user, distilled over time through conversations with many thoughtful contributors in the AAPT and HCDE communities. Attempting to do a review of the literature on the affordances of VR for learning revealed to me that my project is so interdisciplinary as to be overwhelming.

Immersion in another universe with novel physics requires a certain degree of suspension of disbelief to be effective. The immersion of a full 3D virtual environment rather than a stationary 2D window makes an incredible difference in the user's feeling of presence. I understand this difference to be what makes the novel physics an interesting and engaging activity rather than feeling "gimmicky." A 2D simulation would just be dots moving around on a screen. Perhaps students would get into that, but I am skeptical based on my experience of teaching students with PhETs.

Natural 3D interactions in a simulated space allow us as instructors to explore the intersection of the affordances of hands-on labs and those of simulations (e.g., PhETs). Tools are expected to be imperfect in the real world, but emulating such imperfections in a simulator would be considered either a bug or an unwelcome, purely artificial limitation. Using VR lets us create plausible sources of uncertainty in a simulation-based learning tool.

Further, the care and caution students need to take with hands-on apparatus is largely unnecessary in VR; having a dozen 10cm hard spheres flying around one's head would be a disaster in the classroom, but is entirely harmless (and even entertaining) in VR. In this way, using VR creates a digital shield allowing students to express creativity and take risks in ways they would not with hands-on apparatus.

Embodied cognition is a field of cognitive science that focuses on sensorimotor interactions between a learner and their environment as a cornerstone of cognitive development. If interactions driven by one's sensorimotor capabilities are fundamental to the development of cognitive capacities, it stands to reason that a fully immersive VR lab might help students learn better than a pedagogically equivalent 2D mouse-and-keyboard application. There is an argument to be made that the use of VR over a pedagogically comparable simulation (especially of inherently three-

dimensional phenomena) provides benefits to student learning out of the box by virtue of, e.g., emulating the experience of physically handling the apparatus or enabling more natural and intuitive 3D rotations.

Physics education research comparing equivalent 2D and VR simulations of electrostatics, magnetostatics, and phases of the moon have all shown no inherent benefit to the use of VR [51, 33, 40]. If these inherent benefits do indeed exist, they are negligible or overwhelmed by mitigating factors in the context of physics education.

Is leveraging embodied cognition a possible benefit of using VR? Yes. Is it a sufficient reason to use VR? Generally not. My experience and knowledge of the literature suggest that leveraging embodied cognition on its own does little to guarantee that the use of VR is distinctly beneficial. As much as I might want it to be, VR is no educational silver bullet, as it were.

All this to say, NOMR uses VR because it could not be sensibly done without VR, rather than any expectation that VR technology inherently improves any given educational activity. Just as PhETs run in the browser, NOMR uses VR not for the sake of VR but because it is the medium that makes the most sense to use in support of NOMR's pedagogical goals.

One could imagine a 2D version of NOMR: Either a 3D simulation observed on a 2D monitor manipulated via mouse and keyboard or an entirely 2D simulation. Both have serious drawbacks and trade-offs. It could well be the case that those trade-offs are worth it to create an effective learning environment where there is no correct answer without the expense of VR. Anyone who wished to run a controlled experiment along these lines would have my enthusiastic support.

2.5 Scenario designs

Here I discuss two NOMR scenarios code-named Triforce and Tagalong, as their underlying rules are already in print [7]. Multiple other scenarios exist and will not be discussed here for brevity and to keep the "answers," for lack of a better term, out of print. I first present the user-facing experience of the scenario, followed by an account of how the underlying rules behind each scenario were designed.

The user-facing experience of each scenario is not captured in its entirety in this section, as

doing so could be a book unto itself. Instead, salient examples of student lab reports are used to give a depiction of what the classroom experience of each scenario can look like.

2.5.1 Triforce

Students' exploration of Triforce almost always occurs in two phases: Qualitative investigation to determine the number of unique particle types and a scheme for identifying them, and quantitative investigation of the force between one or more pairs thereof as a function of distance.

We will follow the explorations of the Elves, a group of students who investigated the Triforce scenario. Their full report is included in Appendix A.

Their Methods section captures the first phase of exploration, beginning with mashing the "New Particle" button a bunch to observe large-scale emergent behaviors:

> " When we first entered the pocket, we started by spawning many particles to see if we noticed anything substantial in the first minute inside the pocket. We quickly noticed that there were 2 blobs of particles that attracted particles of the same type but repelled particles of the opposite blob. We also noticed a third type of particle that was repelled from itself and didn't have any other particles attached to it. We played around with this third particle and noticed it wasn't either attracted or repelled from one blob. We colored the particles in this blob red. "

So they have one blob they've colored red. All other particles are colored yellow at this point. Some are in another blob, and some are freely floating.

> " We brought [the third particle type] near the [not red] blob and observed it would attach to parts of the blob but sometimes was shot off if we tried it near other parts. We realized that the reason it was getting repelled was that the repulsion force between particles of it's type was most likely stronger than the attraction force between the other particles in the blob that repelled from red particles, something to test later on. "

The Elves have determined that this "third" particle type is indeed one of at least three: It ignores the red blob, is attracted by the core of the not-red blob, and is repelled by other particles of its type. Thus, from observing emergent behaviors, the Elveshave determined that there are three particle types in the scenario.[6] This information in hand, they proceeded to more methodically catalogue the interactions between each pair of particle types:

> " We deleted all the particles and restarted, this time spawning in one particle at a time and labeling them with colors based on if they were attracted to each other or repelled. We went slower this time, checking the relationship between all particles before labeling them. We labeled the particle that was repelled from itself as blue. We labeled the blob that was neutral with the blue particles as red, and we labeled the other blob that attracted green particles as Blue. "

Some groups do a better job of articulating this process than others, but very nearly all groups investigating Triforce end up doing these pairwise comparisons to flesh out their particles' identities. The final result of this process often takes the form of a table as in Figure 2.8, but the Elves took a different approach, as shown in Figure 2.2.

The second phase of investigating Triforce is the quantitative development of a relationship(s) between particle separation and force. The Elvesdecided to take measurements at three different separations for every interaction pair. Groups usually follow a similar process to collect a little bit of information about all interaction pairs, or focus exclusively on the interactions of a single pair such that they can collect more comprehensive F versus r data for that pair. The Elves' force versus distance data for all interactions is shown in Figure 2.3; they hypothesized that force and separation were related by $F \propto \frac{1}{r^2}$ for all interaction pairs, and linearized their plot accordingly to create Figure 2.4.

The Elves concluded:

---

[6] In the Exotic Matter and Manifold Labs, students are told that there are at most three distinct types of particles in their scenario. I always seek to minimize "knowledge from the powers who be" about scenarios, but absent this cap, students will sometimes become unproductively overwhelmed.

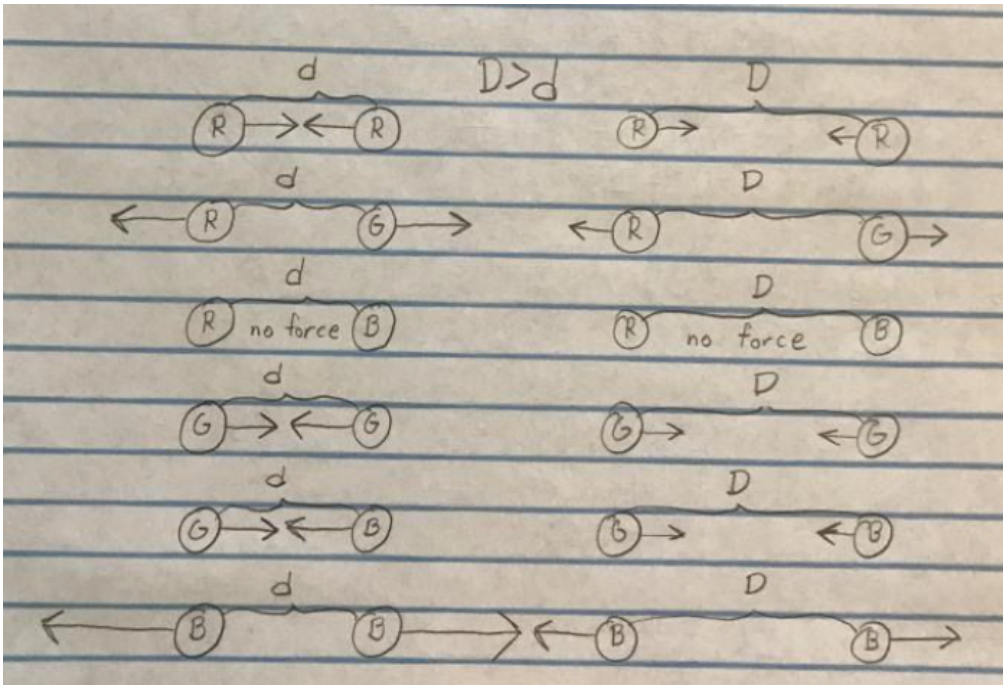


Figure 2.2: The Elves' visual summary presentation of the interactions between each pair of particle types in the Triforce scenario. The left column represents the same interactions as the left, but at a smaller separation. Their three particles, labeled Red, Green, and Blue, were found to interact as follows: Red attracts Red, repels Green, and ignores Blue; Green repels Red, attracts Green, and attracts Blue; and Blue ignores Red, attracts Green, and repels itself.

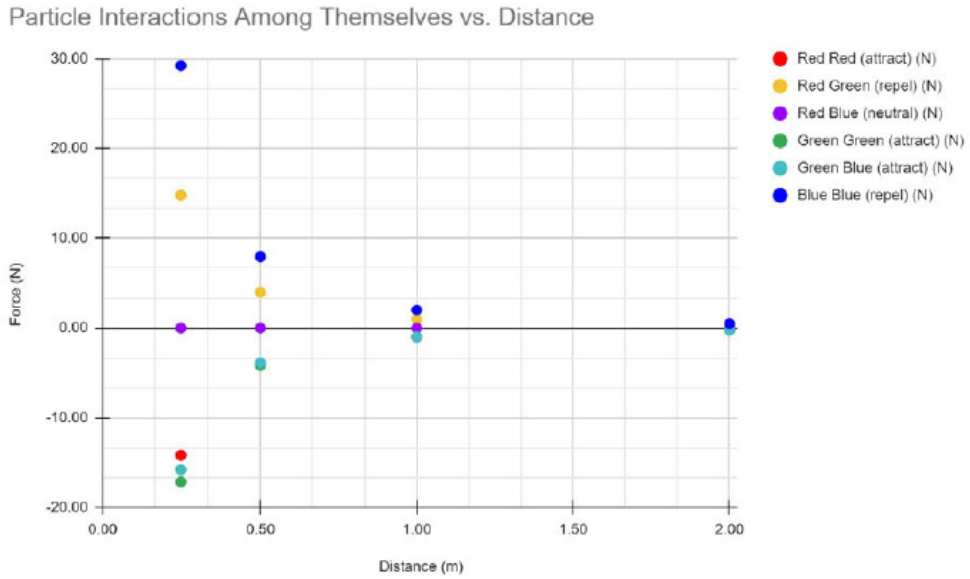


Figure 2.3: The Elves' force versus separation data for all six Triforce particle interaction pairs.

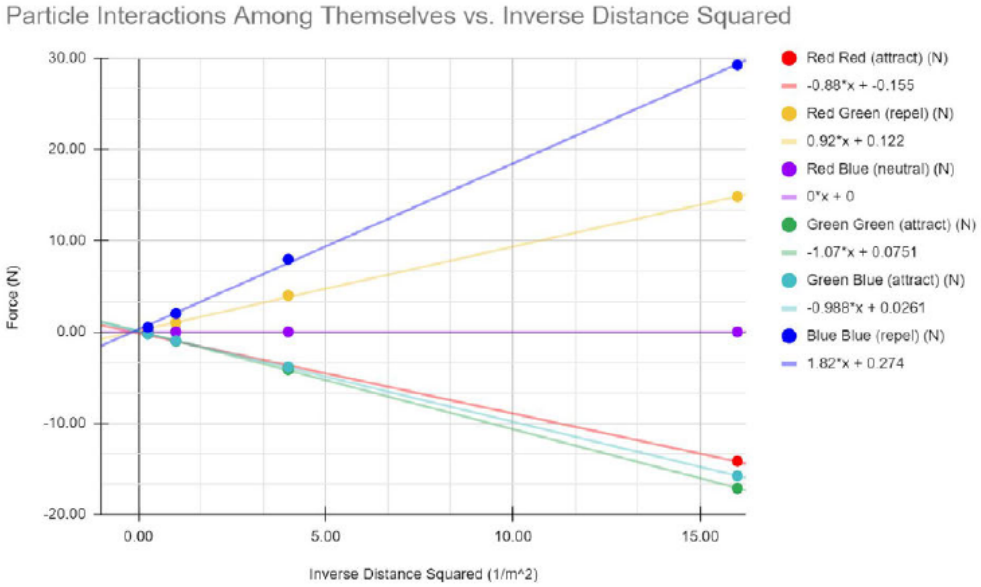


Figure 2.4: The Elves' force versus separation data for all six Triforce particle interaction pairs, linearized by plotting $\frac{1}{r^2}$ on the x-axis..

> " Some of these forces are attractive, and some are repulsive, but the magnitude of the force for all six interactions, including the Red-Blue interaction where there doesn't appear to be a force present, can be described by one equation. Force (N)= $k/(x^2)$, where x is the distance (m), and k is the constant of proportionality with units N(m^2). From the data we collected, we found that k can either be 0, 1 or 2 N(m^2). The value of k is equal to 0 for the Red-Blue interaction, 2 for the Blue-Blue interaction, and 1 for the other four interactions. "

2.5.2 Tagalong

The Tagalong scenario[7] consists of a single particle of the same name. On first glance, the particles seem to *tag along* with each other.[8] One student group dubbed them "sheepy particles" for their tendency to remain in a flock.

The Tagalong scenario is to this day my favorite NOMR scenario on account of the incredible variety of experiments students come up with to investigate it. It is the only scenario for which I will accept a completely qualitative report; I will let the students' descriptions of Tagalongs' behavior in one such report paint the picture. The report is included in its entirety in Appendix A.

We will refer to this student group as the Maples . Their report referred to Tagalongs as "Monke particles" and consisted of a series of qualitative experiments.

Their first impressions gave rise to their first experiment:

---

[7]The code to load the Tagalong scenario in NOMR is 1976, the year Tagalongs first hit the market as part of Girl Scouts' repertoire.

[8]I am very funny.

> “ The first observation we made with regards to the interactions between Monke particles was that their position vectors relative to one another remained constant under steady state; a system of isolated Monke particles experiencing no external forces had all other Monke particles maintain a constant fixed position relative to one another in space. This also means that if one Monke particle was moved, all other particles would ’follow’ it in space. . . . As such, **our first hypothesis governing the interaction between two Monke particles was that the force exerted on one Monke particle was proportional to its velocity relative to the other particle.** ”

In other words, starting with a system of Tagalongs at rest, picking one up and giving it a little toss causes all other Tagalongs to begin accelerating in the same direction, while the tossed Tagalong slows down until all Tagalongs are all moving with the same velocity. If the grabbed Tagalong is not tossed but instead moved to a new position and brought to a stop, all free Tagalongs will come to rest at the same relative position to the grabbed Tagalong as when its motion began.

Students generally alight on one of two ideas from initial observations:

1. The particles all want to stay at the same distance from each other.
2. The particles all want to be moving at the same velocity.

The Maples noted both behaviors but elected to pursue the latter in their first experiment, dubbed the “Relative Velocity Experiment”:

> “ In order to test this initial hypothesis, we conducted a test using the equipment provided to determine the relationship between the relative velocity of two Monke particles and the forces they exert on one another. In this test, the independent variable was the relative velocity of one Monke particle with respect to another, and the dependent variable is the force exerted on the anchored particle. ”

In brief, the students created two Tagalongs, A and B. A was held on a force meter anchored in space, and B was moved away from A, along a line parallel to the force meter. They had

approximate control over the speed of B, and varied that speed to see if the force measured on A varied with it. Their findings formed the core of their model:

> " What we observed through this experiment was that there was a positive correlation between the relative velocity of Monke particles and the force they exerted on each other; faster moving Monke particles seemed to exert more force.... From these observations, we can conclude two things:
>
> 1. That force and velocity were positively correlated in Monke interactions, and
> 2. That the force acting on a Monke particle is in the same direction as other particles' velocities relative to itself.
>
> If any experiment conducted in this lab were to be designated the 'most important', this would be it, as it highlights the single defining factor of Monke particles, in that the relationship between their interactions is determined by their relative velocities. All other experiments conducted were following this observation, and thus this takeaway is the most integral when reporting upon Monke interactions. "

A defining feature of students' interactions with the Tagalong scenario is that a commonly-chosen independent variable, velocity, is not readily measurable within the VLE. There are tools for quantitative distance and force measurements, but none for velocity. This limitation is both a catalyst for a diversity of creative experiments and a frequently-mentioned roadblock:

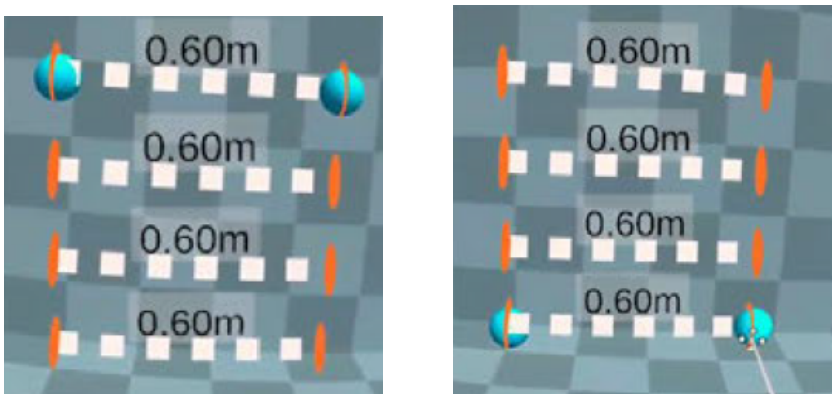
> " We believe that the most significant limitation to this experiment was the inability to measure the speed of the moving Monke particles. As such, for future experiments, we encourage researchers to invest in ways to both measure and control the speed of moving Monke particles in order to properly obtain numerical results for their independent variables. "

The same experimental procedure, with slightly different data collection technique, was used to determine whether the force between two Tagalongs depended on their separation. Rather than examine the force exerted on A while B was moved at varying velocity, the velocity of B was kept

as close to constant as possible, and the force on A was watched as B moved farther away. The force on A did not appreciably change while B was in motion between $\sim 0.5 - 5$ m away, leading the Maples to conclude:

> " **The Monke force acts with the same magnitude irrespective of distance; the only variable found that changes the Monke force is the velocity of one particle relative to another.** "

The Maples also explicitly investigated the direction of the force between two particles to illustrate that force acts parallel to relative velocity rather than, e.g., along a line connecting two particles. This experiment is described in Figure 2.5.



(a) Ruler experiment pre-motion (b) Ruler experiment post-motion

Figure 2.5: The initial and final states of the Maples 's Ruler experiment are shown. The blue spheres are Tagalongs. The rulers were arranged as closely as possible into a plane. The rightmost Tagalong was moved from the top ruler to the bottom ruler, and the leftmost Tagalong's final position was observed. Regardless of the initial distance between the two particles (and thus the angle between their relative position and relative velocity once motion began), here 0.60 m, the left particle moved the same distance as the right particle.

It is common for students to hold on to the idea that all forces are central forces, as they have never seen anything else in context of particle interactions. Listening in on groups beginning to

explore Tagalongs, one will frequently hear students speak of Tagalongs' interactions in terms of attraction and repulsion, trying to frame the behavior they are observing in terms of central forces. After some time, this notion fails to hold up to experimental scrutiny consistently enough that students look to other explanations. As the Maples found:

> " **The Monke force is a non-central force; the direction of the force on a Monke particle is in the direction of the velocity of all other Monke particles relative to itself.** "

Explicitly acknowledging the non-centrality of the Tagalongs' interactions is not common. More frequently, groups implicitly state this finding in their discussion of relative motion being parallel to acceleration.

Next, the Maples investigated superposition of Tagalong forces. Recognizing that Tagalong forces superpose is key to a number of quantitative Tagalong experiments I have seen students devise. If one assumes linear superposition of Tagalong forces, one can infer that if many Tagalongs are anchored in the space, one is moved a specific distance, and one is left free to respond to that movement, the free Tagalong will move a shorter distance proportional to the number of anchored Tagalongs. My first PERC paper contains an early study of a group doing exactly this [7].

The Maples designed two experiments, examining both constructive and destructive superposition. They began by moving a Tagalong parallel to the orientation of a force meter on which another Tagalong was attached; from previous experiments, this gives a roughly constant force reading during the motion of the controlled Tagalong. For their constructive superposition experiment, they added a second controlled Tagalong moving in concert with the first. To examine destructive superposition, they moved one Tagalong toward the force meter and one away from it simultaneously, at the same speed. If force is proportional to relative velocity, and forces superpose linearly, then it follows that they should see double the force in the constructive experiment, and zero force in the destructive experiment. These experiments are depicted in Figure 2.6.

Superposition seemed to hold. In their words:

> " **Conclusion 4: The Monke force acts with field superposition; forces additively and destructively superpose.** "

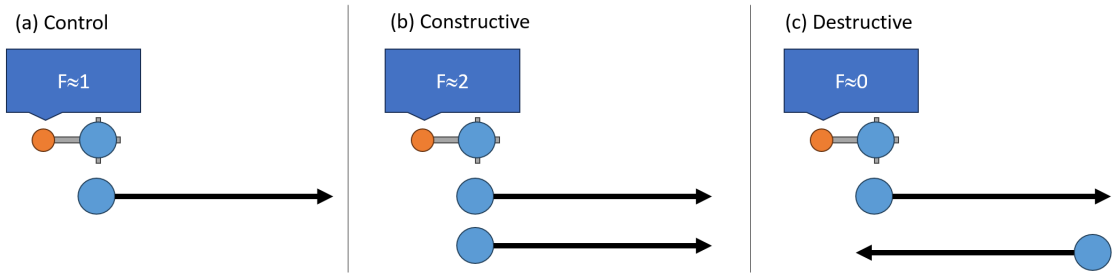


Figure 2.6: Diagrams of the superposition experiments carried out by the Maples are shown. In all cases, a Tagalong is attached to a stationary force meter, and another Tagalong is moved away from that force meter in a line parallel to its orientation. In case (a), that is all that happens, and the force meter read some positive value of force. In case (b), two Tagalongs were moved simultaneously; the force exerted on the stationary particle turned out to be roughly double that of case (a). In case (c), one Tagalong began far away from the force meter, and was moved toward the force meter as the other was simultaneously moved away; during this motion, the force meter read close to zero.

Note that they brought up the idea of a field, which they had not mentioned up to this point. They equate field superposition to the constructive and destructive superposition of forces. It is unclear whether they mentioned a field because their model posits that Tagalong forces are mediated by a field, or because of some reason related to their existing physics knowledge, e.g., they knew the word “field” to be used in the same breath as “superposition” often enough that they felt a need to include it.

They do not mention the idea of a field again, but their final Wave Propagation Experiment lends some credence to the idea that their model was informed by the idea of a vector field mediating the force. They had noted that there is some delay between moving a Tagalong and seeing other Tagalongs accelerate in response. Ordinarily, this would be attributable to inertia; just because a force is applied does not mean an object will instantly come up to speed. However, the Maples were cognizant of this fact, and noted that when a Tagalong was quickly accelerated to some velocity relative to another Tagalong fixed to a force meter, the force read out from the meter continued to increase for a small period of time after the moving Tagalong had hit its maximum velocity.

The Maples devised the Wave Propagation experiment to test whether this apparent delay was a function of distance, implying that Tagalongs' force was mediated by some signal that took time to propagate over distance. Their experimental setups are shown in Figure 2.7.

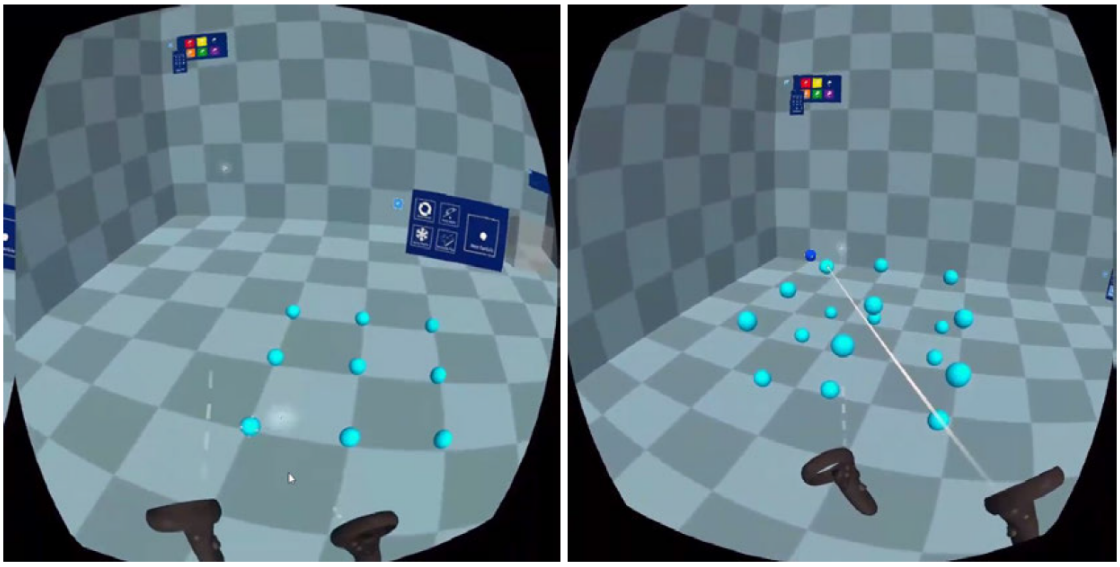


Figure 2.7: The Maples tested for evidence of wave propagation in two and three dimensions using the experimental setups shown.

Many free Tagalongs were arranged in a regular grid; first in two dimensions, then three. One Tagalong from the corner of the grid was picked up and moved a specific distance. If the velocity-force delay was a function of distance, one would expect to see some non-uniformity in the response of the rest of the Tagalongs in the grid, growing in proportion to their distance from the manipulated Tagalong.

The Maples did not observe any such uniformity, and rejected the wave propagation hypothesis:

> " **Conclusion 5: Monke particles do not act on a distance-based delay. By eye, the time it takes for the interaction to occur between a pair of Monke particles is the same irrespective of distance.** "

They noted that it is possible that there is a distance-based delay too small to be captured by their experiment. They maintained that there is some delay, but it is not a function of distance.

The presence of this delay is subtle and not often noticed by students. In fact, it was not an intended behavior in the Tagalong scenario's design; see section 2.5.3 for details about how it came to be. For all that this behavior was not part of the original design, it is still part of the observable behavior of Tagalongs, and thus the modeling thereof is as much fair game as the velocity-force relationship itself.

The Maples summarize their findings thus:

> " All the data that was obtained was exclusively qualitative in nature; no concrete relationships nor mathematical models were derived as a result of the lack of numerical data obtained, though it is speculated that the force experienced by a Monke particle is directly proportional to the relative velocity between it and the Monke particle exerting the force. Thus, future experimentation that takes place must be conducted with the goal of quantitative verification of the observations we have made in our research. For this, we encourage future researchers to obtain equipment that enables them to determine the velocity of objects moving in the Pocket of the Manifold in order to develop well-set dependent and independent variables for quantitative experimentation.
>
> **List of conclusions drawn:**
>
> 1. The force acting on two Monke particles is positively correlated with their relative velocities: faster-moving particles exert forces greater in magnitude.
> 2. The Monke force is independent of Monke particle separation.
> 3. The Monke force is not a central force; the direction of the force exerted on one particle is determined only by the velocity of other moving Monke particles relative to itself.

> 4. The Monke force is cumulative; superposition of the force does exist and works in a summative interaction.
> 5. The Monke force has no time propagation delay; by eye, all Monke particles experience the same force at the same time in the presence of another moving particle, irrespective of distance between them.
>
> Finally, we felt as though a note regarding the naming of the Monke particle should be included as an endnote to this report. Based on our initial observations, we determined that the particles followed each other, and off of this, we determined that the 'Monke particle' would be an appropriate name for such a particle, as "monkey see, monkey do". "

The Maples claimed they were unable to measure velocity due to lack of equipment; many groups get around this limitation by using a stopwatch or screen recording in conjunction with the measuring tape tool. However, those workarounds have their own limitations, and as the Maples showed, there is a great deal of fruitful investigation to be done without velocity as a quantitative independent variable.

2.5.3 Development of underlying rules

NOMR's physics are all field-mediated interactions between pairs of point particles. The classic examples of this type of interaction are gravitational attraction and electrostatic attraction and repulsion. To begin developing fictitious phenomena within this umbrella, let us dissect these familiar laws of physics and think about where we can riff on them.

Consider Coulomb's law for the force $\vec{F}$ between two point particles with charges q_1 and q_2 separated by displacement $\vec{r}$, with magnitude $r = \|\vec{r}\|$ and direction $\hat{r} = \vec{r}/r$:

$$\vec{F} = \frac{kq_1q_2}{r^2}\hat{r} \tag{2.1}$$

To start, we list the features of this relationship to help figure out which ones we can take creative liberties with:

- The *Coulomb constant* k is a constant universal scaling factor.

- *Charges* q_1 and q_2 are positive or negative scalars intrinsic to the interacting objects, where negative value represents negative charge, and positive value represents positive charge.
- The roles of the direction $\hat{r}$ and magnitude r of the displacement vector $\vec{r}$ can be considered separately:
  - *Direction* $\hat{r}$ determines the direction of and is always parallel to $\vec{F}$.
  - $\vec{F}$ falls off with the square of *distance* r.

The Coulomb constant k actually has to be varied to give us something playable without having to juggle unwieldy values of mass and charge. Two 1 C charges a meter apart will exert force on each other comparable to the weight of the Great Pyramid of Giza. Ultimately, I ended up setting $k = 1$ Nm^{2}C^{-2}, $m = 1$ kg for all particles.[9], and tuning charge magnitudes q_1 and q_2 in the range of $|q| = 1 \sim 3$ C.[10]

Let us separate the r-dependence out into $f(r)$. Starting with $\vec{F} = kq_1q_2f(r)\hat{r}$, we can say $f(r) = 1/r^2$ to get Coulomb's law. There is nothing stopping us from making $f(r)$ any well-behaved function of r. This fact is the basis of several NOMR scenarios. However, the interactions between each pair of particle types in a scenario must share the same $f(r)$, as I discovered in the development of the Triforce scenario.

Triforce

Electric interactions occur between two types of charge: positive and negative. What if there were three types of charge? Exploring this question gave rise to the Triforce scenario.[11] I began by looking into the strong nuclear force, which acts between three colors of quarks: red, green, and blue. I quickly realized that the interaction between two quarks cannot be tidily described in tidy

[9]Relevant: https://xkcd.com/2205/. It is worth noting that this value of k, keeping the same value of magnetic permeability of free space, implies that the speed of light in the NOMR universe is about 3,000 m/s. This particular implication is not reflected in NOMR's physics code.

[10]This turns out to be helpful both for the developer's understanding and the VLE's stability; real-time physics engines start to run into issues with very tiny or very large (non-infinite) masses.

[11]For the uninitiated, the scenario's name and the names of its particles reference the Legend of Zelda series of video games. The code to load the Triforce scenario in NOMR is 1986, the year the original Legend of Zelda game was released for the Nintendo Entertainment System.

force functions. This way lay madness, so instead I decided to riff on electrostatics.

We can write a more general force law for the interaction between two particle types i and j: $\vec{F}_{ij}(\vec{r}) = A_{ij} f_{ij}(r)\hat{r}$, where $f_{ij}(r)$ encapsulates the distance dependence of the force law governing the interactions between particle types i and j. Let us consider only interactions where force $\vec{F}_{ij}$ is parallel to separation $\vec{r}$, letting us drop the vector notation: $F_{ij}(r) = A_{ij} f_{ij}(r)$.

In an interaction that has only a single type of charge (i.e., gravity), there is only one interaction pair to consider: two objects with that lone charge type. Electrostatic interactions following Coulomb's law occur between two types of charges, giving rise to three interaction pairs: positive-positive, negative-negative, and positive-negative. The former two are repulsive, and the latter is attractive; this is captured in Coulomb's law by representing charge as a signed scalar quantity, where positive charge is represented as a positive number, and negative charge as a negative number.

Then an interaction acting on three types of charge, call them A, B, and C, would have six interaction pairs: A-A, A-B, A-C, B-B, B-C, and C-C. The force laws between each of these interaction pairs must all individually satisfy Newton's 3rd law, but they need not all be the same force law. How different can they be? What can we alter? We have two options: The constant A_{ij} and the r-dependence $f_{ij}(r)$. We know the former can be varied between interaction pairs thanks to electrostatics: A is negative (attractive) for interactions between oppositely-charged particles and positive (repulsive) for interactions between like-charged particles.

This leaves the question of whether f_{ij} can vary for different interaction pairs. By Newton's third law, we must have $f_{ij} = f_{ji}$. Let us imagine two types of fictitious particles α and β. Two types of "charge" gives us three interactions: $F_{\alpha\alpha}$, $F_{\alpha\beta} = F_{\beta\alpha}$, and $F_{\beta\beta}$.

At this point we must acknowledge a particular constraint of the exotic matter physics (EMP) engine I created to drive NOMR's physics, documented in the Process chapter. The force upon a particle is determined from the value of one or more vector field value(s) evaluated at the particle's position, and intrinsic properties of the particle. Force laws are not defined between pairs of particle types: Instead, each particle has one or more rules dictating the global vector field(s) it generates, and one or more rules dictating how it reacts to the field(s) at its location. In other words, we need

to be able to write our force law for each interaction pair as a combination of two expressions: A field generated by one particle $E_i(r)$, and the force exerted by that field on the second particle $F_{ij} = E_i(r)A_j$.

Let us break the scaling constant A_{ij} into components intrinsic to each particle: $A_{ij} = A_iA_j$. Breaking $F_{\alpha\alpha}$ and $F_{\alpha\beta}$ down into field generation and field reaction expressions gives us the following, where E_i represents the field generated by particle i:

$$F_{\alpha\alpha}(r) = A_\alpha A_\alpha f_{\alpha\alpha}(r) \quad \rightarrow \quad E_\alpha = A_\alpha f_{\alpha\alpha}(r), \;\; F_{\alpha\alpha}(r) = E_\alpha A_\alpha;$$
$$F_{\alpha\beta}(r) = A_\alpha A_\beta f_{\alpha\beta}(r) \quad \rightarrow \quad E_\alpha = A_\alpha f_{\alpha\beta}(r), \;\; F_{\alpha\beta}(r) = E_\alpha A_\beta.$$

Those expressions for E_α suggest $f_{\alpha\alpha}(r) = f_{\alpha\beta}(r)$, meaning we cannot have different r-dependence for each interaction pair. It is worth noting that this limitation is due to assumptions made in the EMP engine rather than violation of physical laws. If one were to implement the engine using pairwise interactions rather than monopoles reacting to fields, I do not see any reason why one could not have multiple r-dependence functions in a scenario. However, any such phenomena would not be consistent with a field theory analogous to electromagnetic or gravitational fields, which may have unforeseen implications for, e.g., students' concept of field superposition.

So what does this mean for the Triforce scenario? To create a scenario with three charge types, we can vary the interaction pairs' proportionality constants A_{ij} such that each particle type can be identified by whether it attracts, repels, or ignores a second particle of each type. This is summarized in Figure 2.8.

As for the r-dependence, we have concluded that it has to be the same for all interactions, so we can just pick whichever $f(r)$ makes for interesting behavior. I ended up going with the classic $\frac{1}{r^2}$ after toying with a couple of alternatives.

This choice of proportionality constants gives rise to some interesting emergent behavior: If one throws a bunch of all three particle types into the space in roughly equal proportion, they tend to form two clumps, or condensates. Without reading on, can you guess the composition of the two condensates from the information in Figure 2.8?

	Power	Wisdom	Courage
Power	-1	1	-1
Wisdom	1	-1	0
Courage	-1	0	1

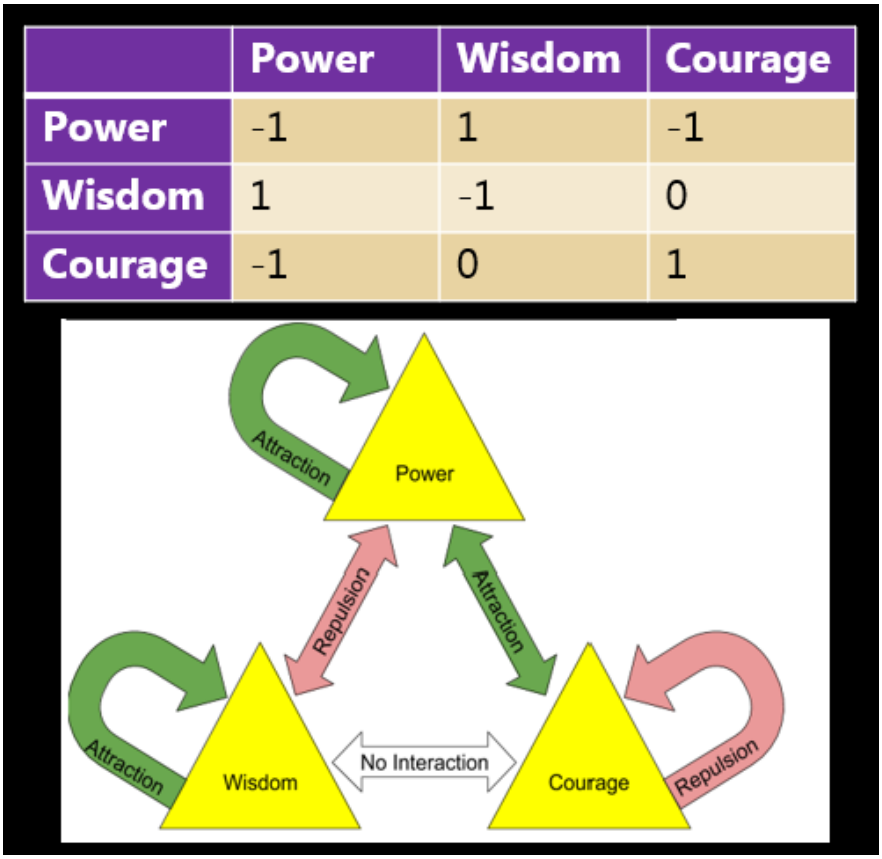


Figure 2.8: The Triforce scenario's particles and their interactions are visually represented. The table at top represents the proportionality constants A_{ij} for each interaction pair. These values give rise to the interactions depicted below: Each yellow triangle represents a distinct particle type. Negative proportionality constants cause attraction; positive ones cause repulsion. Setting a proportionality constant to zero causes the particles to effectively ignore one another.

Power and Wisdom repel each other and both attract themselves, so they are in different condensates that want to avoid each other. Courage does not interact with Wisdom, is attracted to Power, and repels itself. On its own, Courage would not form a condensate. However, one ends up with a shell of energetic Courage particles around the condensate of Power particles. The very strong attractive force of a clump of Power particles allows the Courage particles to overcome their repulsion to each other, for certain ratios of the number of Power and Courage particles. I hold out hope that one day, a student group will tell me what that ratio is, and if it varies with number of particles beyond $n \sim 10$.

Tagalong

In gravity and electrostatics, force is always parallel to displacement. What would it look like if force was not parallel to displacement? This question gave rise to the Tagalong scenario.

I wanted to create a force law dependent on some variable other than displacement. My first thought was to riff on the magnetic term of the Lorentz force $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$. This quickly proved problematic: The magnetic term depends on the cross product of two extrinsic vector variables. Every other scenario's force law(s) depend only on changes in a single extrinsic vector variable: separation. Coming in with no foreknowledge of magnetic interactions, even figuring out which variables are relevant would be a major challenge, let alone setting up controlled experiments in which only one variable is changed.

We could swap out an extrinsic vector quantity with an intrinsic one–say, a dipole moment, which would be easier to control. Alas, the orientation of a dipole moment would itself comprise an extrinsic variable, leaving us with no less complexity than we started with. I came to the conclusion that forces dependent on more than one vector variable were best put aside for a difficulty level more suited for professional physicists.[12]

So if we can only have force depend on a single vector variable, that rules out any function using a cross product. The force function needs to be given direction somehow. An early idea

[12] https://xkcd.com/2207/

to make force only act along, say, the x-axis, was tossed because it implies a universal reference frame.

This constrains the design space to direct proportionality between force and relative velocity $\vec{F} \propto \vec{v_{rel}}$, which on first glance seemed too simple to constitute a meaningful scenario. This case in particular highlights the importance of remembering that it is not the complexity of the rules that matters, but the complexity of the behaviors they give rise to. I created the Tagalong scenario with a single particle type and a simple force law: The force between any two particles opposes and is proportional to their relative velocity. $\vec{F} = -A\vec{v_{rel}}$, where A is a positive constant tuned to taste. Absent a velocity meter, this gives rise to an incredible breadth of creative experiments.

Limitations of the physics engine also contributed to Tagalongs' user-facing behavior. There are multiple different ways to measure the velocity of an object in Unity.[13] All of them are very noisy and often disagree with each other. The noise made for a force that was all over the place and occasionally produced strange behavior. The cleanest solution I could find at the time was to average each particle's velocity over the last n physics engine updates, and use that average velocity for the purpose of determining the force between Tagalongs. Tuning to taste, I set $n \sim 20$, which reduced the noise to the point it was not obvious while not creating so much lag between one particle being moved and another accelerating in response.

It is worth noting that Tagalongs violate physical consistency Rule (2) of scenario design: Angular momentum is not conserved. It took me a while to realize this infraction, and we had already used Tagalongs for many groups of students by the time it hit me. If one considers Tagalongs in context of orbital motion, the plausibility of their existence implodes into a puff of logic pretty quickly. However, they do a tidy job of conserving linear momentum. I have deliberately not thought too hard about whether energy is conserved in the hope that an enterprising group of students will answer that question for me one day.

---

[13] One can pull velocity from changes in the position of an object as recorded by its Transform either every physics frame or every display frame, or from the velocity field of its Rigidbody every physics frame.

2.6 Instructor preparation and classroom contexts

2.6.1 Role of the instructor

The most straightforward way to describe the role of the instructor in NOMR labs is as an advisor. Graduate students' advisors do not know what the outcome of their mentees' research will be. Nonetheless they serve to provide guidance and help students focus their work in such a way as to move their research forward in productive directions. NOMR instructors' role is very much the same: Without regard for any specific end result, help students develop and execute on their research questions. In interacting with students during model-generating NOMR labs, it is of paramount importance to first understand the students' current observations and understanding of the phenomenon, and strictly build off of that basis.

It is here that the importance of never sharing the "correct answers" with TAs comes into play. Speaking from experience, it is all too easy to unconsciously guide students toward an answer one has come to expect. I firmly believe that if I could erase my memory of the NOMR scenarios, it would make me a more effective instructor for NOMR activities. Just as it takes a great deal of effort to bring students away from a confirmatory mindset, many instructors are uncomfortable not knowing what they are "supposed to" guide students to. The end result of an experiment is the least important element of a model-generating NOMR activity and is never meant to be compared to some source of authority; internalizing this idea is often as much a challenge for instructors as it is for students. Just like the students, strictly keeping the "answers" under wraps helps to make this transition easier.

Instead, it is recommended that instructor preparation focus on being familiar with the in-headset behavior of the phenomena in question and developing their own toolbox for exploring the phenomena.

To boil down the most important points covered in NOMR instructor preparation for model-generating experiments:

- The correct answer is the one that students can make the best case for.
  - A satisfactory answer does not need to be a complete model of the scenario in question,

nor does it even need to be accurate. It must only be supported by the students' data and analysis and be testable.

  - Different groups will get different models for the same scenario, and that is fine.
  - Frequently remind students that there *is no right answer.* The majority of students' previous experience in science labs is one exercise in confirmation after another. This paradigm shift can be distressing for some students.
- Developing students' sense of ownership and agency is critical.
  - Frequently remind students that this is *their* experiment. They are researchers exploring an undocumented physical phenomenon.
  - Encourage students to name the particles they discover.
  - Avoid pushing students along a railroad.
  - Encourage groups to pursue wild ideas. Students pursuing creative questions and experiments you have not seen or thought of before is a sign that NOMR is working exactly as intended.

Example student-instructor interaction

Suppose a group of students is investigating and developing a mathematical model for minty particles, which repel each other at close distances and attract each other at distances greater than some threshold. A group of students is capturing force vs. distance data for a pair of minty particles with distance as the independent variable. They have decided to take data only for values of distance greater than the turnaround distance. In this case, an appropriate reaction would be to ask the group, "How did you select the values of distance that you will measure force at?" Determine whether this is a deliberate experimental choice, a limitation of their methodology, or simply an oversight.

1. The group may answer that they decided to aim for a very precise model of the attractive behavior, to exclusion of the rest of the minty particles' behavior. In this case, the independent variable value choices are deliberate and well-motivated, and the group should not be dissuaded from that choice.

2. It may be that the group could not figure out how to measure force when the particles were repelling, or were confused about how to reflect attractive and repulsive forces in their data. In this case, help students resolve the difficulty in question, and suggest that taking data in the repulsive region would help them develop a more complete model.
3. It is not uncommon for a group of students to re-use the same values of distance from a previous NOMR lab, i.e., a quantitative model-testing experiment on Coulomb's law. In this case, the TA should prompt students to think about the significance of choosing certain values of distance. It is helpful to make a comparison between electric and minty interactions: Though electric interactions only have one distinct region of interest, is that true for minty particles?

2.6.2 Classroom structure

Three general principles of instructional design have particular implications for how NOMR is designed to be used in the classroom: Accessibility, affordability, and supporting teamwork.

Accessibility means that all students must be able to participate in NOMR labs–notably including students who, for any reason, do not use VR. There must be meaningful, substantive roles in NOMR activities that can be completed in their entirety without ever putting on a headset. VR technology is far from being universally accessible. Headsets are generally not well-designed to permit a user to comfortably wear glasses while using them. Many students are simply not interested in using a VR headset; susceptibility to motion sickness is a common reason.

Affordability goes without saying and is mentioned for completeness. Nobody has an infinite budget. 1:1 headset-to-student implementation of VR is inordinately expensive even after the dramatic reductions in VR hardware pricing over the last few years.

Supporting teamwork is a goal of the entire introductory physics lab curriculum at UW, and this should be reflected in NOMR's design as well. Science is a team sport, and teaching students to be physicists includes teaching them to develop ideas and produce deliverables as a cohesive, functional team.

NOMR is built from the ground up as a team activity. **We have found that groups of 2-**

4 students work best. Each group uses a single VR headset. The headset user's view is streamed to another display visible to the rest of their group. In this way, only one student per group ever needs to use a VR headset (though turn-taking is strongly encouraged among students who are willing to use the headset). Further, no activity-specific instructions are included within the VR headset, nor is it possible to directly record data within the VLE. Thus, other members of the group must verbally guide the headset user in accordance with the lab manual and aid them in taking data.[14] It is common to see a group of four students adopting distinct roles as the headset user, their handler, the data scribe, and the data plotter.

Using one headset per group reduces the cost of outfitting the classroom with VR headsets by a factor of 2-4, while at the same time working to ensure that no students are left behind by the use of VR and creating a need for effective teamwork.

For specific details about the choice of VR hardware and screen mirroring setup used by NOMR, see Chapter 3.

2.7 Examples of lab activities using NOMR

NOMR is designed such that the VLE itself is not tied to any specific curriculum or instructional activities. In this section I give a description of how NOMR is used in 100-level and 200-level lab courses as of the time of writing. The complete lab handouts are included in Appendix B. I again thank Kazumi Tolich for her thorough iteration on the 100-level NOMR lab activities at UW. I wrote the first version of the lab handouts as a prototype for a wild research idea, and she brought them to maturity.

This content is also covered, albeit more briefly, in Chapter 4 due to its origin as a standalone paper. I go into more depth here.

I use the same codes to refer to the lab activities in each case.

Table 2.1 gives examples of 100-level and 200-level lab curricula at UW in which NOMR plays a major role.

---

[14] This format works just as well in person as remote; see the process chapter.

Wk	**100-level**			**200-level**		
	Lab	**Name**	**Type**	**Lab**	**Name**	**Type**
1	0	Coulomb's Law	T	E1	Electron Beam pt 1	T
2	VR1	Charge	T	E1	Electron Beam pt 2	C
3	VR2	Minty	G	N1	Nuclear Decay pt 1	T
4	VR3	Exotic Matter	G	N2	Nuclear Decay pt 2	C
5	B1	Bulbs pt 1	G	VR1+2	Charge + Minty	T / G
6	B2	Bulbs pt 2	T	VR3	Exotic Matter	G
7	B3	Capacitors	G	VR4	Manifold Proposal	C
8	(No lab)					
9	B4	Unknown Resistor	A	VR5	Manifold Testing	T
10	(No lab)			VR6	Final Presentations	C

T = Testing, G = Generating, A = Application, C = Communication

Table 2.1: This timeline of events shows the curriculum 100-level and 200-level lab students complete over the course of the quarter. The labels **G**, **T**, **A**, and **C** represent model **generation** experiments, model **testing** experiments, **application** experiments, and **communication**, respectively.

2.7.1 Preparatory activities: Charge and Mint

See Appendix B for an example Charged Particles Lab (VR1) handout used in a 100-level course.

Model-generating NOMR lab activities throw a lot of new ideas at students. Rather than throwing them directly into the deep end, we ease them in with a model-testing experiment and a generously-scaffolded model-generating experiment. The 100-level course completes these activities in two separate weeks; the 200-level labs combine them into a single week.

The model-testing experiment (VR1) has students design and conduct a qualitative test of Coulomb's law as it exists in NOMR. The purpose of this experiment is twofold: (1) Familiarize students with the VLE, and (2) develop the experimental design and data analysis skills students will need to create mathematical models from scratch in future labs. It is expected that Coulomb's law is quite well known to students at this point; it is used as a familiar context in which to become acquainted with a new technology and experimental "apparatus," as it were. Specific learning goals include:

1. Learn to use the VR headset and NOMR VLE.
2. Understand the force and distance measurement tools in the VLE.
3. Design an experiment to test a known model with clearly-defined independent, dependent, and controlled variables.
4. Quantify the uncertainty associated with taking measurements of force and distance in NOMR.
5. Linearize data according to a model and create a best-fit line.
6. Use measured uncertainties to make a judgement about a model's fit to empirical data.

The final product of VR1 is a plot of force versus distance data, linearized by plotting force F against one over the square of distance $1/r^2$; an example is shown in Figure 2.9. Students use this plot to judge whether the trendline is a good fit and thus whether Coulomb's law should be rejected or not.

See Appendix B for an example Minty Particle Lab (VR2) handout used in a 100-level course.

The model-generating experiment (VR2) exposes students a NOMR scenario composed of a single particle type: the minty particle. Minty particles repel each other at close distance and

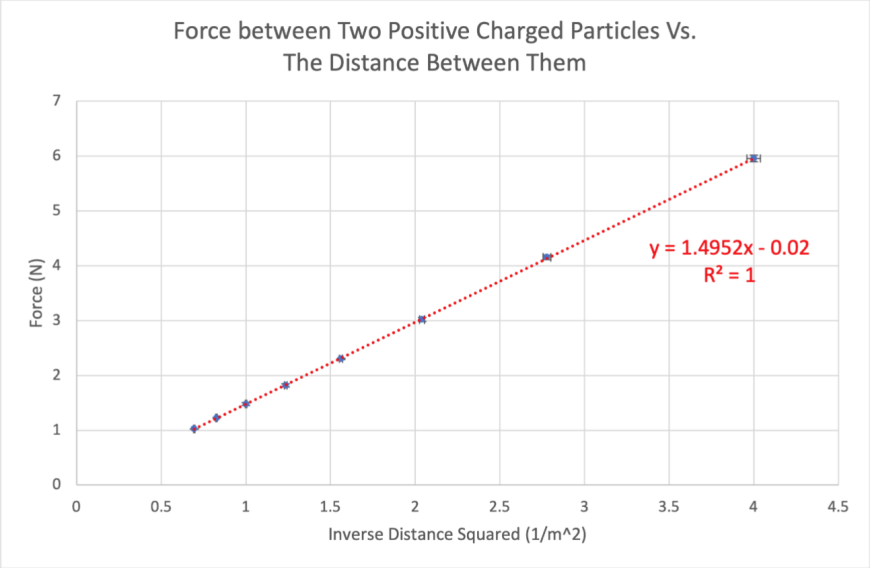


Figure 2.9: An example of a linearized plot of force versus distance data for a pair of charged particles, produced by a group of students completing the Charge lab (VR1).

repel each other beyond some distance threshold. Students' investigations of minty particles are scaffolded by making frequent analogies to their investigation of Coulombic interactions in VR1. VR2 includes scenario-specific scaffolding in order to ease students' transition into a science lab environment where there is no correct answer to confirm and they are the ultimate authority on what the "answer" should be. Students are tasked with developing a *testable* empirical model for the interactions between minty particles. The model need only be testable and supported by their data; it need not be comprehensive or even accurate.

A key process introduced to students in VR2 is the use of qualitative observations to inform the development of an empirical model. Examples of observations of the behavior of two minty particles and their implications follow:

- Minty particles repel when they are near each other, and attract when they are far away. A turnaround point where the force is zero exists at a certain separation between particles. Thus, a model relating force and distance must have an x-intercept.

- If two minty particles are brought as close as possible to one another and released from rest, they appear to undergo oscillatory motion. Looking at the full range of distances achieved in this motion, the range of distances for which the force between the particles is repulsive seems to be shorter than the range of distances for which it is attractive. This could imply that for separations where force is repulsive, the magnitude of force scales up faster than in the range of separations where the force is attractive.
- Beyond the repulsive region, no matter how far apart two minty particles are moved, they continue to exert upon each other a substantial attractive force which increases with distance. Thus, a model relating force and separation cannot go to zero at large distances like the Coulomb force does–at least on the scale of the size of the virtual workspace.

Students are not immediately given access to quantitative force and distance measurement tools. Instead, they begin by making qualitative observations such as those above and developing a preliminary model based on those observations. Once students have done so, they are given a code that unlocks quantitative tools. Then they design and conduct an experiment to develop an empirical mathematical model for minty particles' interactions, using many of the tools they developed in VR1. In particular, they use various linearizations to try out possible models for the separation dependence: For example, if they think the force might be proportional to one over the separation, $F \propto \frac{1}{r}$, they plot F versus $1/r$ and determine if that produces something acceptably close to a line.

2.7.2 Exotic Matter Lab

See Appendix B for an example Exotic Matter Lab (VR3) handout. The Exotic Matter Lab is only called that in the 100-level labs; in 200-level labs, it serves as the first week of the Manifold Lab, described in Section 2.7.3.

With VR1 and VR2 under their belts, students are prepared to tackle VR3, in which each group is given one of a handful of different scenarios to investigate and generate a model from. Unlike VR2, there is no scenario-specific scaffolding; every group receives the same lab manual, even though each group is working with different physics. The open-endedness and absence of correct answers for which NOMR is designed fully comes into play.

Students are tasked with creating a testable empirical model of interactions between particles in their scenario. The key words are *testable* and *empirical*: Their model must be backed by qualitative and/or quantitative evidence, and it must be experimentally testable and falsifiable. Their model need not be comprehensive, nor accurate, nor strictly focused on finding force as a function of distance. Students are well and truly on their own to design, conduct, and analyze an experiment to create a brand new scientific model.

Instructors hold the codes used to enable each scenario in the NOMR VLE. At the beginning of class, they distribute one code to each group in the section. This allows for some amount of differentiated instruction; some scenarios are simpler than others, and the nature of the challenge varies between scenarios. Based on the instructor's experience with their students, they can assign more complex scenarios to high-functioning groups and simpler scenarios to students who struggled with VR1 and/or VR2.

2.7.3 Manifold Lab

See Appendix B for example Manifold Lab (VR3-6) handouts used in a 200-level course.

The Manifold Lab is a three-week activity followed by a presentation week, It is designed to take students through a complete cycle of the scientific process, generating and validating new knowledge as part of a class-wide collaboration. It is framed as something of a role-playing game. From the lab manual's introduction:

"The Manifold is a recent discovery which has granted us entrance into myriad pocket universes, each with its own new forms of matter which obey unique rules for interacting with each other. The four fundamental force laws that govern our universe are known to us, and the force laws of the matter in the Manifold seem to exist alongside them. Its discovery has prompted a scientific gold rush; though no matter can be brought back from the Manifold, it represents a font of knowledge no research scientist can resist. You are researchers tasked with understanding and cataloguing the force laws that govern the Pockets of the Manifold. The National Science Foundation (NSF) supports the exploration of the Manifold and offers grants to promising research groups for cutting-edge measurement tools and expanded Manifold access."

The instructor plays the role of an NSF program manager and journal editor. In the first week (VR3), they provide research teams with pilot grants to access a Pocket of the Manifold (give each group a scenario code). The first week proceeds otherwise just as the Exotic Matter Lab does in 100-level labs: Absent scenario-specific scaffolding, each team develops a testable empirical model from their scenario.

In the second week (VR4), students review each other's work, and then every team's report with their model is passed to a team that did not explore the same scenario. Every team ends up with another team's report documenting a model of a scenario that they have not themselves experienced. Based on the other team's report, each team writes a grant proposal detailing an experiment to test the other team's model. In this proposal, each team has four credits to spend on new and improved equipment they can use to perform different or more precise experiments.

At this stage, students are reminded that confirmation is not the goal of a testing experiment; the possible outcomes of a testing experiment are to reject or fail to reject the model in question. In the nicest, most collaborative way, students are instructed to try to poke holes in the models they were given. I make a habit of referencing high-energy particle physics experiments as examples during this discussion. For one, the muon $g - 2$ experiment (a project UW Physics is heavily involved in) is designed to narrow the error bars on a previous measurement of the muon's magnetic dipole moment. This is interesting because the experimental and theoretical values of that dipole moment are just far enough apart to raise eyebrows, but not enough to reject the Standard Model; the $g - 2$ experiment seeks to whittle those error bars down to a point that the experimental value deviates strongly enough from the Standard Model's prediction that it becomes evidence for the rejection of the Standard Model. All that to say: Testing experiments in experimental physics are purpose-built to reject models, and designing them with that goal is not in any way an attack on the creators of a model; it is just good science.

The instructor (wearing their NSF hat) reviews each team's grant proposal for specificity and feasibility, allowing time for a round of revisions. Once the instructor deems a proposal worthy, it is "funded" by giving the team the codes to unlock the upgraded equipment they requested.

The following week (VR5) sees students carrying out the testing experiments they proposed.

They write up their results and present them in a conference symposium format in the following week (VR6). The teams that generated and tested each model interact at this point to reflect on the process of developing and validating the model. I make it a point to have some discussion questions prepared to kick things off, but I often do not have to use them; dialogue between the teams emerges organically following the presentations.

Chapter 3

IMPLEMENTATION

Chapter 2 explains the what and why of NOMR; this chapter explains the who, when, and how. NOMR has been in active development for over four years. During that time, many different people have contributed to its growth and adoption. This chapter gives a chronological account of the development of the NOMR VLE, the instructional material and classroom implementation practices enabling its use, and the community supporting it.

Sections 3.1 and 3.2 describe the VR physics instruction prototypes I created over Spring 2018 - Summer 2019 and ultimately abandoned before NOMR came to be. In Fall 2019 and Winter 2020 (Section 3.3), the key ideas behind NOMR slot into place, a prototype is developed, and it is pilot tested with a group of honors students. The project's progress is abruptly interrupted by the onset of the COVID pandemic in Spring 2020 (Section 3.4).

In Summer 2020 (Section 3.5), NOMR's interactions engine is overhauled and NOMR is adapted for use in hybrid and online classroom models. The adapted version of NOMR is field tested in hybrid form in Fall 2020 (Section 3.6) and then in online form in Winter 2021 (Section 3.7); during the latter field test, the instructional material used with NOMR is rewritten and expanded. Over Spring and Summer 2021 (Section 3.8) the NOMR VLE is heavily optimized, experimental multiplayer is added, and the measurement tools' UX is overhauled.

In Fall 2021, NOMR becomes a permanent part of the introductory physics lab curriculum at UW, and my knowledge of and responsibility for the practices around classroom implementation and TA preparation for NOMR are gradually transferred to the UW Physics Department's instructional team over the following quarters (Section 3.9). Beginning Summer 2022 and continuing to the present day (Section 3.10), the NOMR Student Developer Group takes over the continued development of the NOMR VLE and my focus turns to the evaluation and documentation of NOMR's

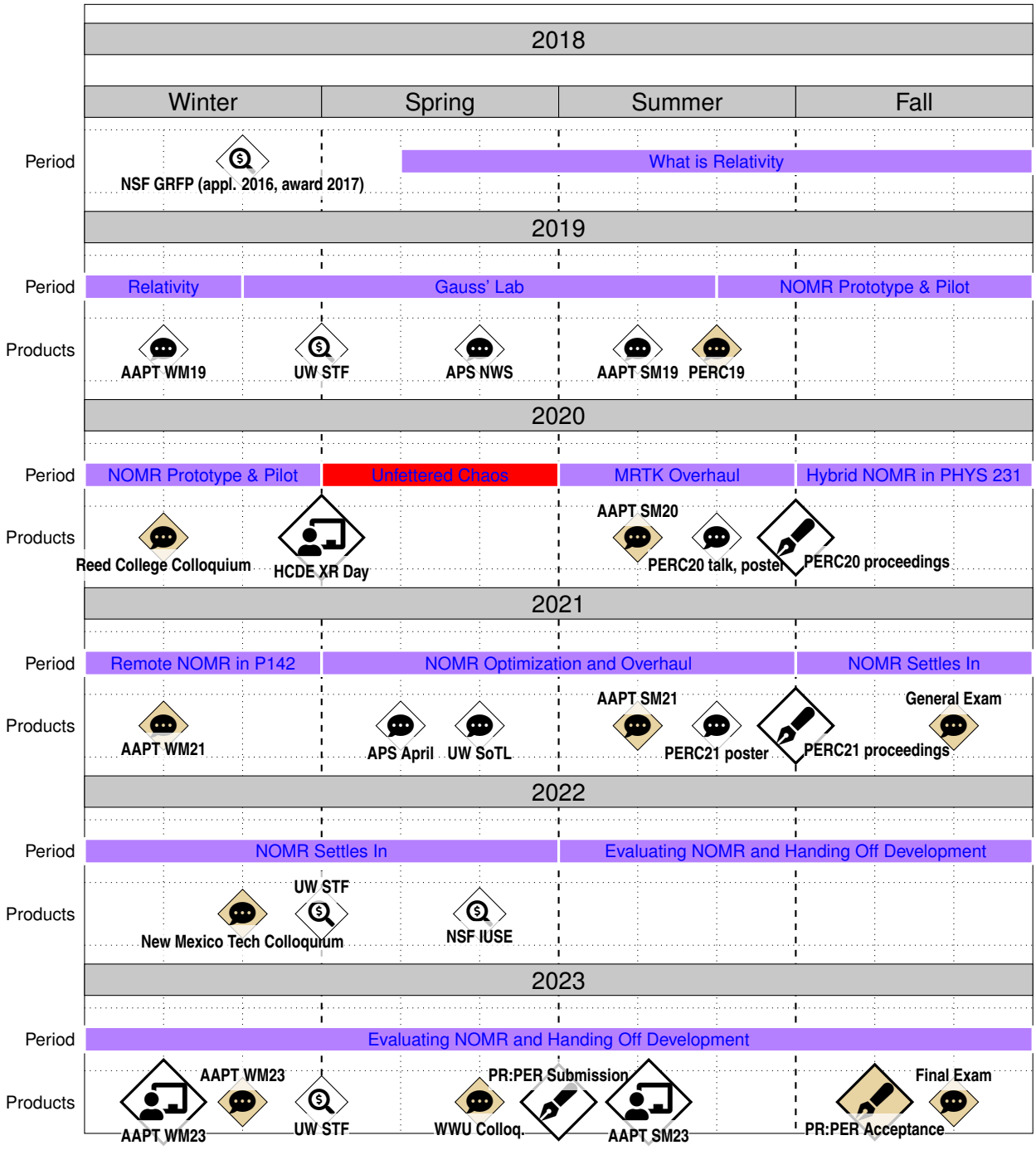


Figure 3.1: A timeline of NOMR and its precursors' history and talks (💬), papers (🖋), workshops (💻), and grant proposals (🔍) produced in connection with them. Gold backgrounds represent invited works.

effects on student learning.

3.1 Spring - Autumn 2018: What is Relativity

This section and the next briefly describe VR physics projects I worked on before NOMR came to be.

I wanted to make a VR special relativity simulator, so anyone could experience a world where the speed of light is comparable to a brisk jog. MIT Media Lab's *A Slower Speed of Light* [30] is a short game where the player wanders around collecting items that slow down the speed of light. As the speed of light decreases, the player is able to see relativistic effects become gradually more apparent. I thought that being able to experience special relativistic effects arising from one's own motion could make for a helpful teaching or outreach application. I adapted the special relativity visualization code published by MIT Media Lab and created What is Relativity for the HTC Vive; see screenshots in Figures 3.2 and 3.3.

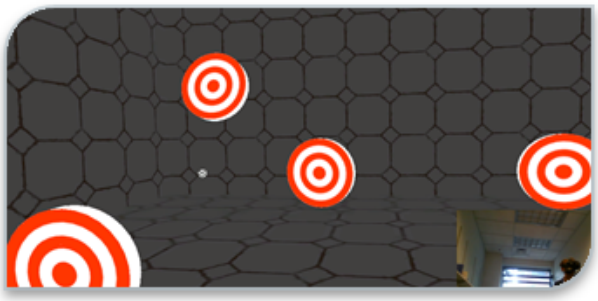


Figure 3.2: A screenshot from What Is Relativity's gallery shooter demo, taken with a very high speed of light setting.

Moving in a universe with a very slow speed of light, length contraction makes the world bend and twist around you as the Doppler effect manifests a violent rainbow of color. This is neat, if a bit overwhelming, when viewed on a 2D monitor. When rendered in immersive 3D, it turns out to be a very effective motion sickness generator. A world that induces motion sickness faster than a dinghy in a storm is not conducive to education.

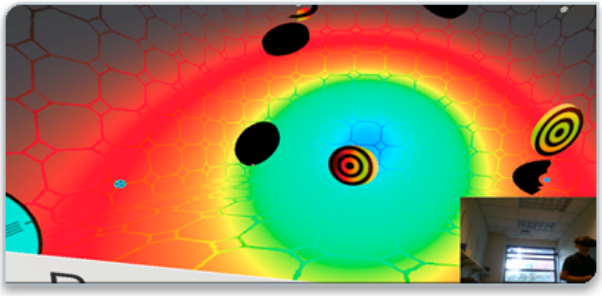


Figure 3.3: A screenshot from What Is Relativity's gallery shooter demo, taken with a very low speed of light setting.

Ultimately, I came to the conclusion that I needed to be more experienced as both a VR developer and as a physics instructor to be able to make What is Relativity into something useful in the classroom. I turned my attention instead to simpler concepts in electromagnetism.

Adin Pierce, a PHYS 499 student, worked with me on the special relativity simulator and some electrostatics visualization from October 2018 to May 2019. He came up with an idea for a data structure to help with managing event causality in context of special relativity; we developed this structure to fruition and it turned out to be quite a boon. I would still like to write it up someday.

(Return to timeline)

3.2 Winter - Summer 2019: Gauss' Lab

3.2.1 Motivation

In coming to understand the UW PEG's work on identifying student misconceptions as a basis from which to build curriculum [36, 47], I found that students had particular difficulties with details around electric flux and Gauss' law. It occurred that a VR application with dynamic field line and electric flux visualizations could potentially help students with these concepts in ways tough to achieve with 2D media.

The resulting application was dubbed Gauss' Lab. It ran on the original Oculus Quest, released in May 2019, just in time for the Gauss' Lab pilot test in a PHYS 122 lab in July 2019. It was

well-received by students and was featured in UW Perspectives.

3.2.2 Technical developments

Electromagnetic simulation and visualization

In 2012, Dr. Jim Pivarski wrote *Magnetodynamics*, a Unity library simulating electric and magnetic interactions for a variety of geometries. Though it was no longer publicly available in 2018, Pivarski kindly granted me a copy of the library over email. *Magnetodynamics* formed the basis of the electromagnetic simulations used in Gauss' Lab.

The package included visualization options in the form of field vector arrows and electric and magnetic field lines. I added to these capabilities by implementing a fragment shader[1] to visualize flux through each point of a 3D object (Figure 3.4) and adapting the field line code to run on the GPU rather than the CPU (Figure 3.5).

Interaction design: Locomotion and object manipulation

VR is unique as a medium in that any VR application must be designed for two environments simultaneously: The VR environment and the user's physical environment. Locomotion is one of the first and most important problems a VR designer must address: How does a user navigate the virtual space? To what extent will they accomplish that by physically moving about their play area? In-headset locomotion that is not reflected in the real world is a common cause of motion sickness; achieving in-headset locomotion by moving about in the real world is a common cause of collisions, trips, and falls.

For classroom usage, it is most safe and convenient to have the VLE be a seated experience. For educational usage, the VLE should be as accessible as possible, meaning that in-headset locomotion that may risk inducing motion sickness is off the table. It follows that the safest solution is the complete absence of locomotion for both the user's physical body and their virtual avatar.

---

[1] A shader is a program that tells the computer how to display an object. A fragment shader is a specific type of shader that operates on a pixel-by-pixel basis.

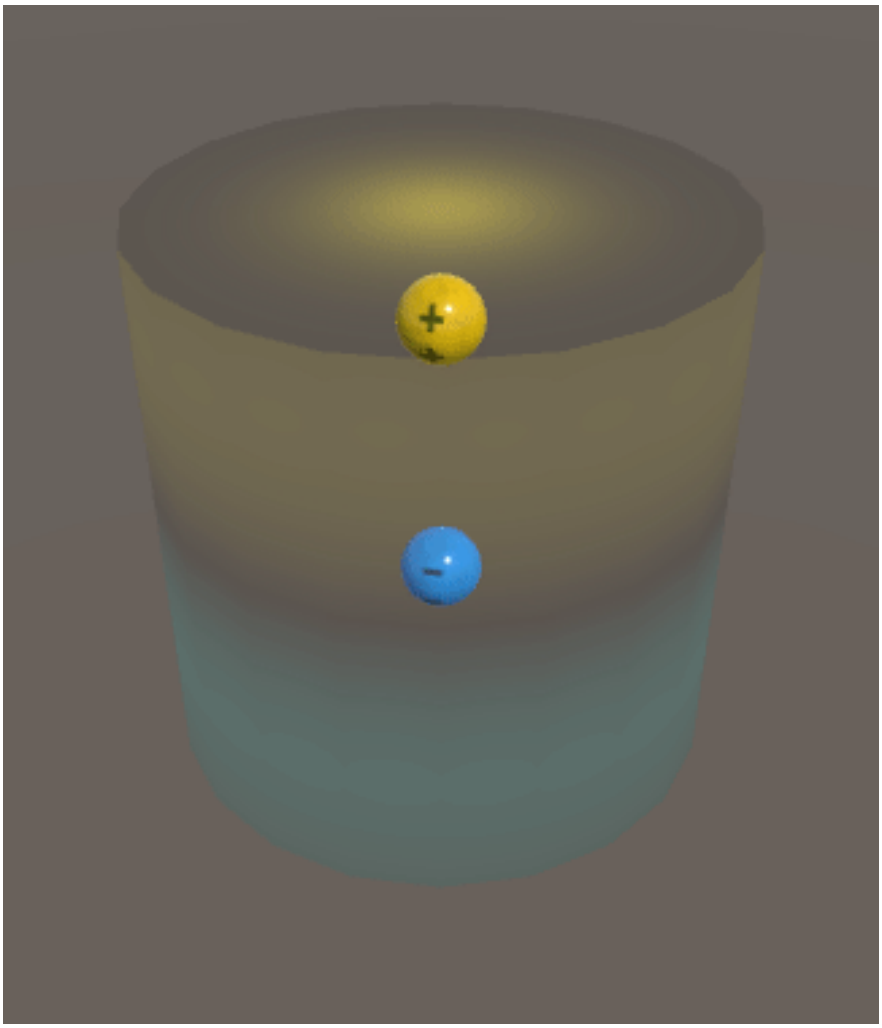


Figure 3.4: A positive and negative charge are separated vertically and placed on the axis of a cylinder. The sign and magnitude of the resulting electric flux through each point on the cylinder is represented by the color and intensity of the shading on the cylinder, respectively.

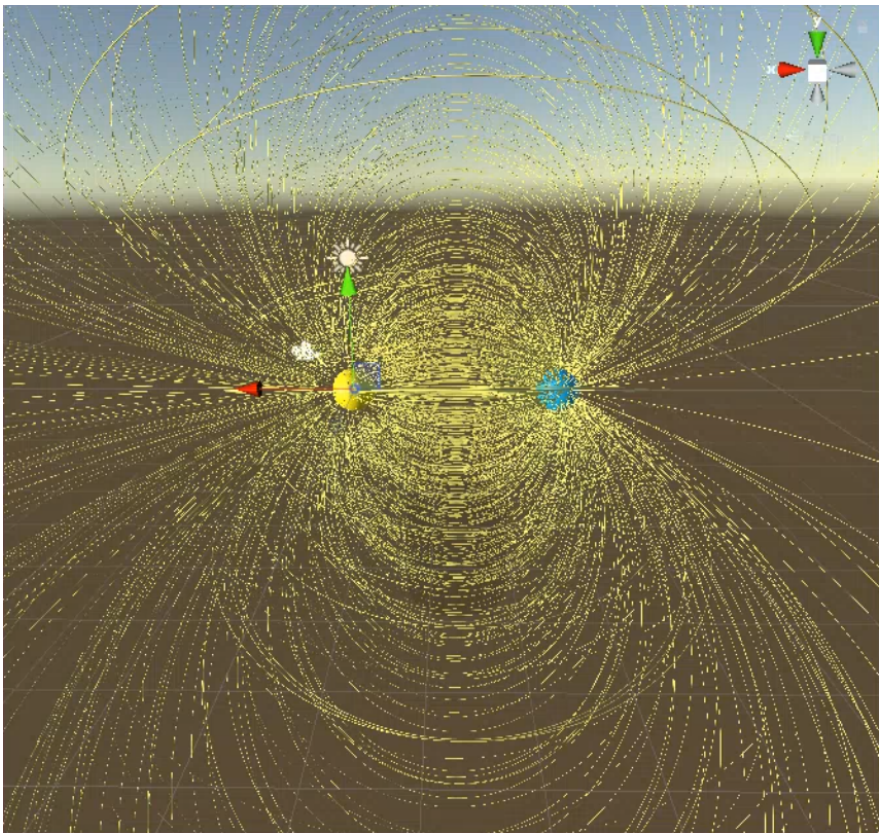


Figure 3.5: The electric field line distribution of a horizontal electric dipole is shown. The distribution consists of ~ 2000 field lines, rendered in real time on the GPU. All field lines begin at the positive charge (left) and terminate at the negative one (right). As the calculation of a field line halts after some large distance, many fail to reach back around to the negative charge, resulting in the asymmetry seen here.

I played a handful of games and experiences on the Quest to understand how existing works enabled users to interact with objects anywhere in a room-sized space without any physical or virtual locomotion. The Bond-inspired VR escape room game *I Expect You to Die* [45] stood out. In every scene, the player's avatar is seated; even if the player chooses to physically stand up and move away, their avatar will not move outside of a small box around their seat. The player must interact with a number of objects within and beyond arm's reach to solve puzzles and progress through the game.

To interact with objects beyond the player's reach, the game uses a remote interaction system flavored as telekinesis. The user points at an object, holds down a button to "grab" it, uses the thumbstick to move the held object toward or away from them, and achieves rotation and other dimensions of translation simply by moving the controller. This process is shown in Figure 3.6.

I replicated this object interaction design pattern using the open-source Virtual Reality Toolkit (VRTK), implementing some custom code to support moving a held object toward and away from the user in response to thumbstick input.

3.2.3 Classroom implementation practices

Due to accessibility concerns and limited availability of headsets, students worked in pairs with one headset to a pair. In this way, any student unable to or uncomfortable with using a VR headset could complete the activity with their partner's assistance.

The lab manual prompted students to pass the headset to their partner at the end of each section to encourage turn-taking. To ensure that both partners were able to meaningfully participate at all times, the headset's display was mirrored onto a desktop PC on the lab bench (Figure 3.7) using `scrcpy` [21].

I developed a set of safety protocols which were included in the lab manual and the pre-lab presentation:

1. Headset safety
    (a) If you are not comfortable using a VR headset, you are not required to do so; you will work with a partner to complete this lab.

Figure 3.6: Four screenshots from Schell Games' *I Expect You to Die* [45] illustrate the "telekinesis" object manipulation design pattern used to allow stationary players to manipulate objects outside of their reach.

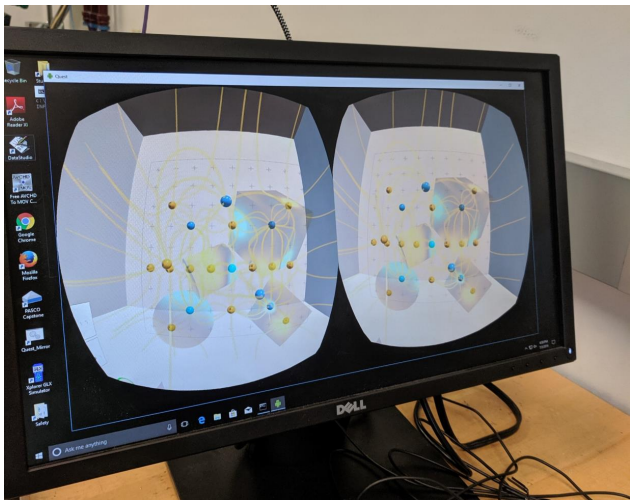


Figure 3.7: The in-headset view of Gauss' Lab is mirrored onto a desktop PC.

Figure 3.8: A motley assortment of graduate students and a summer undergrad pre-pilot test Gauss' Lab and stress-test my ability to facilitate a plurality of headset-using pupils.

(b) Should you experience nausea or similar discomfort while using the headset, take it off immediately and notify an instructor or TA.

(c) Wipe the headset facial interface foam with the provided antibacterial wipes between users. Take care not to wipe the headset lenses.

(d) Use the wrist straps attached to the controllers.

(e) Use the headset only while seated. Do not stand and walk around while using the headset.

2. Observer safety

(a) **Do not touch people in VR without their consent.** If you need to tighten a strap on the headset or help your partner find a button on their controller, ask for permission before touching them.

(b) Be aware of people in headsets around you. Give them space. They cannot see you, and may move unpredictably at any time. Make sure to warn them if you're passing through their space.

(c) Take care not to trip on the cable connecting your headset to your lab PC.

3.2.4 Learnings

As pilot tests of new technologies go, the classroom implementation went well. There were no "game-breaking" bugs or technical difficulties; every group was able to finish the activity. The activity was received positively by students, likely thanks in no small part to the novelty of VR. I observed some playful activities beyond the requirements of the lab manual, both by students and by colleagues (Figure 3.9).

No students reported feeling symptoms of cybersickness, nor did any students fall, get whacked with a controller, etc. due to use of VR. Safety-wise, the pilot test went perfectly.

It took on average 30-45 minutes for each group to familiarize themselves with the VR headset and software to the point that they could begin the lab activity itself.

I used questions on tutorial pretests and final exams to collect pre- and post-intervention data on students' understanding of electric flux and Gauss' law. Comparing to historical data with the

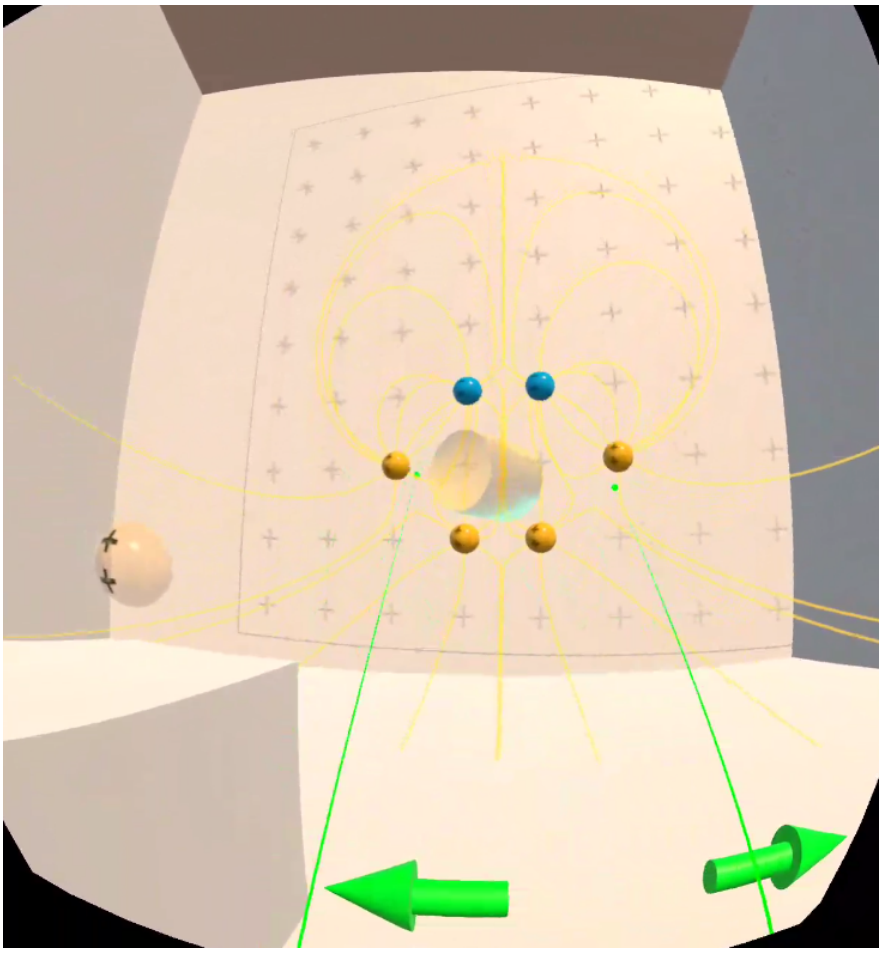


Figure 3.9: A screenshot of Gauss' Lab shows the VLE with a handful of charges placed by one Hannah Sabo in such a way as to create a field line distribution resembling the visage of Squidward, a character from the cartoon Spongebob.

same questions, the pre-post gains were not large enough to be consistent with the hypothesis that Gauss' Lab significantly improved students' conceptual understanding.

In retrospect, Gauss' Lab was most useful in that it spurred the development and pilot testing of classroom VR usage protocols and the VLE interaction design. Both in terms of technical development and my experience facilitating VR in the classroom, Gauss' Lab motivated the creation of the foundation upon which NOMR would be built.

(Return to timeline)

3.3 Autumn 2019 - Winter 2020: NOMR prototype & pilot

As I worked to salvage Gauss' Lab's null result in Autumn 2019, I met with the UW Physics Department Introductory Lab Coordinator, Jack Olsen. In this meeting, he mused: Could VR be used to simulate physical phenomena not present in our universe, to allow students to discover completely new phenomena with no known answer to measure themselves against?

That little comment gave rise to NOMR. I repurposed much of Gauss' Lab's code and assets to develop a prototype of NOMR's VLE, first administered in Winter 2020 in the honors electromagnetism lab. Students were given a set of particles with unknown properties in the VR environment and tasked with developing scientific models describing the particles' interactions.

This section describes development of the prototype VLE's physics and interactions, the original lab manuals, and the practices used to implement VR in the classroom. We conclude with a summary of outcomes and reflections from the pilot test.

3.3.1 Exotic Matter Physics engine

To create the fictitious laws of physics simulated by NOMR, I developed the Exotic Matter Physics (EMP) engine, a generalized field-mediated particle interactions simulator. The engine is "generalized" in that no fields are hard-coded into it; the electric, magnetic, and gravitational fields the budding physicist is familiar with are not first-class citizens, but instead are treated the same way as any other field we may decide to invent.

Field-mediated interactions between EMP objects are defined through the fields an object generates and the fields it reacts to. For instance, to create electrically charged monopoles, one would give an object a scalar charge value and add two EMP behaviors to the object:

1. Generate a field labeled "Electric" that is proportional to this monopole's charge and falls off as $1/r^2$.
2. React to an Electric field at this monopole's location by applying a force equal to the product of the Electric field value and the charge.

To implement a full Lorentz force $\vec{F} = q\left(\vec{E} + \vec{v} \times \vec{B}\right)$, one could add other objects that generate a field with a different label (say, "Magnetic"). Then, to the electrically charged monopoles, one would add a field-reaction behavior that applies a force equal to the cross product of the monopole's velocity and the Magnetic field at its location, multiplied by the monopole's charge.

EMP is written in C# for Unity. I reproduce the GitHub documentation for EMP's five core classes below:

`InteractionField`: This object represents an instance of a field that can mediate forces between particles, which is the basis of EMP. It has only two pieces of information associated with it:

- `name`: Anything you want to call it, so long as there are no two fields with the same name. Name is used to tell `InteractionFields` apart.
- `visualColor`: Used for visualizations where only one color is appropriate, e.g. editor gizmos and field sensors.

One would use different `InteractionFields` when there is a need for different force laws to coexist in the same scene. That will make sense once we go over generators and reactors.

Note that unlike all other classes in EMP, this class inherits from Unity's `ScriptableObject` rather than from `MonoBehaviour`; this means that it can exist independently from `GameObjects` in Unity scenes, as its own asset on our filesystem that can be reused across contexts.

`EMStrength` This abstract class represents the magnitude of some form of "charge" that exists

on an object. This allows EMP to add support for various different types of charged objects (monopoles, dipoles, quadrupoles, global uniform fields, etc) with relative ease. However, for the moment, the `MonopoleStrength` implementation is the only one used in NOMR labs.

`FieldGenerator` This abstract class contains all information about the direction and magnitude of the field produced by an electromagnetic object, by implementing the abstract method `FieldValue`, which takes a world-space position and returns the vector value of the field at that position due to the object with which the `FieldGenerator` is associated.

`FieldReactor` This abstract class contains all information about a charged object's response to an `InteractionField` at its location. By implementing the abstract method `ForceUponReactor`, subclasses of `FieldReactor` can take a 3D vector representing the direction and magnitude of a specific `InteractionField` at its location and calculate the force experienced by its associated object due to that `InteractionField`.

`EMController` `EMController` makes the magic happen. This class inherits from `Singleton`, which gives it an additional feature on top of `MonoBehaviour`: It provides a single means of access (`EMController.Instance`) which can be called from any other script, will create an `EMController` if one does not already exist, and will ensure that no more than one exists at any given time. Every `FieldGenerator` and `FieldReactor` automatically registers itself with `EMController` on creation. Singletons are considered in some circles to be a cursed game programming pattern, but I haven't found a way around it that isn't an order of magnitude more complicated.

Every frame, `EMController` tells every `FieldReactor` in the scene the value of its `InteractionField` at its position (looping over all FieldGenerators of the same field type to do so), and tells the `FieldReactor` to apply that force to its `Rigidbody` component. All of that happens automatically; there are no properties exposed for `EMController`, and if you need to go modifying something in `EMController`, chances are that you may be barking up the wrong tree (or something has gone terribly wrong).

3.3.2 Prototype VLE

The Gauss' Lab VLE served as the starting point for building the NOMR VLE. The same interactions engine (VRTK) was employed and the design of the workspace itself remained very similar, save for the addition of a virtual workbench designed to mirror the location and dimensions of the user's physical lab bench.

All forms of field visualization were removed.[2] The electromagnetics scripts employed in Gauss' Lab were swapped out for EMP and two new tools were added: The force meter and the measuring tape.

Particle anchors and pausing physics

How one anchors particles in space was also overhauled. In Gauss' Lab, there was a pre-set 2D grid of particle anchors that could be rotated about the vertical axis by increments of (I think) $30°$. Pre-pilot prototypes of the NOMR VLE included anchors as a distinct tool, which could be spawned and placed freely by the user. Any particle brought close to an unoccupied anchor by the user would snap to the anchor when released. Initial out-of-classroom testing proved this approach to be cumbersome. One TA remarked that they would like to be able to just "click" the particle and have it freeze in place. In response to this feedback, anchors were removed as their own tool. Instead, particles were configured such that aiming a pointer at a particle and clicking the trigger button toggles between anchored and unanchored states for that particle.

A global "Pause Physics" button was added, which freezes all moving objects in the VLE. This option made configuring experiments much less cumbersome to the user. Unlike anchoring individual particles, pausing physics does not zero out the velocity of a particle, instead caching the velocity from before physics was paused. When physics is unpaused, particles resume moving at their pre-freeze velocity. However, moving or anchoring a particle while physics is paused will

[2] I spent over a month on the GPU field lines. Scrapping all that work stung. Achieving physically accurate 3D field line distributions for arbitrary configurations of point charges turns out to be a very complex problem. For starters, all field lines emerging from a monopole should be equally spaced. There is no way to place n equidistant points on a sphere for most n, and cranking n up until this issue stops mattering only produces more problems. Field lines are a mess and a half and I encourage anyone following in my footsteps to not.

cause its cached velocity value to be set to zero.

Force meter and measuring tape

The force and distance measurement tools provide quantitative data about particles' interactions. They must strike a balance between maintaining the uncertain or "imperfect" nature of their real-world equivalents, and being easy to use in the VLE.

The design of the force meter began with (and ultimately remained very much alike) a real-world spring scale. Spring scales use a spring with a known constant to translate the spring's stretch into a force reading. This reading captures the component of force exerted parallel to the orientation of the spring scale. For example, holding a spring scale horizontally and hanging a weight from it will yield a reading of zero. One rarely encounters this situation in the usual course of introductory physics labs; a spring scale in such a context will nigh universally be oriented parallel to the direction of the force it is measuring.

This behavior is translated verbatim into NOMR's force meters, excluding the physical stretching of the spring. In NOMR, the force exerted on the object attached to the force meter can change direction rapidly in space and in time; if the force meter were to stretch in the course of taking a measurement, this stretch would cause the attached object to move to a different location than where the user intended to take their measurement. In some situations, the object could end up oscillating back and forth on the force meter. Thus, physical stretching of the force meter was not included.

In absence of lab benches and stands on which to attach force meters, the NOMR force meter has an "anchor" end which remains fixed in space at all times unless moved by the user. The other end is the "grabber" to which an object can be attached; the grabber and the rod connecting it to the anchor will rotate freely about the anchor in response to forces upon the attached object.

The NOMR force meter displays a scalar equal to the dot product of the force $\vec{F}$ currently exerted on the object attached to it and the orientation $\hat{r}$ of the force meter, a unit vector pointing from the force meter's anchor to the grabber to which the object is attached: $F = \vec{F} \cdot \hat{r}$. The force is displayed in Newtons with two decimal places of precision.

Now, consider measurement of distance. All NOMR interactions are designed to be achievable with a single controller, for accessibility and to minimize cognitive load. The force and distance measuring tools in the VLE are no exception. Imagine using a tape measure with one hand and no assistance. Some adjustments were necessary to make this a non-frustrating experience in VR.

I ultimately made the ends of each measuring tape independently movable. That is, a measuring tape cannot be picked up in its entirety with one controller; instead, each of its two ends must be individually moved to the desired positions. The distance between these two points (in meters) is displayed on the tool's readout, with precision down to the nearest cm. I occasionally hear minor dissatisfaction with this approach, but none of those grumbles have included a better idea, so the measuring tape works as such to this day.

3.3.3 Prototype instructional material

The lab handouts created for the pilot implementation of NOMR was co-written by myself and Jack Olsen with input from Prof. White Brahmia. They followed the design principles of other transformed lab manuals developed in that quarter:

- Less is more; a lab manual should be at most 2 pages long.
- Align with ISLE; each experiment should be cleanly recognizable as a model testing or model generating experiment.
- Students should be developing and testing models from quantitative and qualitative empirical evidence.
- Leave the details of the experimental process up to the students as much as is feasible.

The resulting handout was short enough that I include it here:

Lab 5

Model Generation

Your goal this week is to create a model describing the interactions between never-before-seen particles in a VR lab.

Using the VR Lab

Your group's VR lab comes with one to three distinct kinds of particles, which are generated at random each time the "New Particle" button is pressed. Every group has access to a different set of particles.

- All particles will look the same on creation. You can use the new Paint tool (see instructions) to color them if you choose to do that.

Build Your Model

Use the table on the following page to document your model as you work through today's lab.

Start up "Cassia: Lab 5 V*X*" (where X is a number from 1 to 4) in the headset and the desktop mirror and begin to explore the particles in the VR lab. Develop observational experiments to answer the following questions:

- ***How many unique types of particles does the VR lab generate?***
- ***How can the unique physical properties of a particle be determined?*** *(recall from the first lab this quarter that you explored two types of charge, and deduced that the same types attracted each other and different types repelled each other)*

Document the setup and results of each experiment. Come up with names by which to refer to each particle. Once you have identified your particle(s), develop and conduct experiments to characterize their interactions. Your results should answer the following questions:

- ***What independent variables affect the force between a pair of particles?***
    - ***What quality/qualities of your systems affect the force between a pair of particles?***
- ***What is the functional relationship of each independent variable to the force between particles?***

Develop a mathematical model for the interactions between particles that is consistent with your answers to the above questions and supported by qualitative and/or quantitative observational evidence.

Apply Your Model

If time remains, use your model to answer the following optional questions.

- *What structures do the particles form, and how does your model explain their formation?*
- *Do the particles' interactions satisfy conservation of momentum and energy? Does your model?*
- *Does your model suggest that particles' interactions are mediated by fields? What fields do your particles generate? How can you express your model in terms of these fields?*

<table>
<tr><th></th><th>Charged Particles (example)</th><th>Your Particles</th></tr>
<tr><td>Particle Interactions</td><td>
<table>
<tr><td></td><td>Yellow</td><td>Purple</td></tr>
<tr><td>Yellow</td><td>Repulsion</td><td>Attraction</td></tr>
<tr><td>Purple</td><td>Attraction</td><td>Repulsion</td></tr>
</table>
</td><td></td></tr>
<tr><td>Particle Identities & Properties</td><td>Positron: A particle with charge q = +1. We painted these yellow.
Electron: A particle with charge q = -1. We painted these purple.</td><td></td></tr>
<tr><td>Independent Variables</td><td>Charges q_1 and q_2 of the two particles; these can only have the values +1 or -1. Distance between particles r, and direction of the displacement vector between them $\hat{r}$.</td><td></td></tr>
<tr><td>Interactions: Symbolic Representation</td><td>$\vec{F} = C\frac{q_1 q_2}{r^2}\hat{r}$</td><td></td></tr>
<tr><td>Interactions: Descriptive Representation</td><td>The force between two particles with charges q_1 and q_2 is proportional to each charge and a constant C. The magnitude of force decreases with the square of the distance between the charges r and acts in the direction of the displacement vector between them $\hat{r}$.</td><td></td></tr>
</table>

3.3.4 VR classroom implementation

In the transformed lab series, students worked together in groups of 3-4 for the entire term. Gauss' Lab used one headset per pair of students, and we thought to keep a similar ratio in this case by providing each group with two headsets. We were concerned that one headset in a group of four students would create difficulties in keeping the entire group engaged at once. Ultimately, we found that most groups only used one headset and ignored the other. In the few exceptions, groups used the two headsets simultaneously to either get the entire group familiar with the in-headset experience quickly, or to collect data faster.

3.3.5 Learnings

Pedagogy: The NOMR pilot blew us away. Students exhibited immense creativity in ways none of us had seen before in our experiences as instructors. The Tagalong scenario proved one of the more difficult scenarios but produced the most novel and creative experiments. Students were able to generate testable mathematical models from all scenarios.

Being able to paint particles to label them was essential in scenarios with more than one particle type.

Classroom management: The activities were completed satisfactorily within the two-hour lab period. TA intervention was in ample supply due to the presence of three "additional TAs" (myself, Jack, and Suzanne), but it seemed feasible that the class could be handled by 1-2 TAs familiar with the NOMR software and troubleshooting the screen mirror software. No VR-related safety incidents occurred.

Technical: The screen mirror software cuts out randomly and sometimes will not start for no apparent reason. It is by far the weakest link of this entire setup.[3]

Many little UX problems persist from Gauss' Lab. The menu buttons occasionally do not

[3]This is still the case in 2023, though it has improved somewhat by buying cables with better-fitting USB-C connections and using a script that auto-restarts scrcpy when a headset is disconnected and reconnected. To learn more about classroom management of VR headsets, contact the UW Introductory Physics Lab staff, David Aplin (daplin@uw.edu) and Eddie Mendoza (eddiemen@uw.edu).

respond. The “tractor beam” push/pull function while holding an object gets pretty dicey when an object is brought too close to the controller. Tool readouts can overlap each other and become hard to read, and there is no way to position them independent of their tools.

Particles can get stuck outside the virtual lab box and continue to interact with particles within it, leading to unexpected behavior. They can be removed with the “reset scene” button, but the particle rigidbody physics should be re-visited to prevent clipping outside of the box, and killzones should be added outside of the box to delete any particles that do make it through.

Students can pick menus up to arrange them as is most useful to them. This is usually helpful, except when a student drags a menu outside of the lab box and lets go, leaving it stranded and inaccessible. The only way to get it back is to restart the app.

(Return to timeline)

3.4 Spring 2020: Unfettered chaos

Friday March 6, 2020 was the last day of the pilot test of NOMR. It was also the last day that UW courses were held in person until Fall 2021. The following few months were utter disarray; I made little progress on my own work, instead helping with the emergency virtualization of the physics lab curriculum.

Undergraduate independent study student Hayden Potter worked with me during this time. He explored a couple of ways to simulate conductors in Unity. We found that restricting an ensemble of charged particles to the surface of a conductor by overriding the position of each particle to the closest point on the conductor’s surface each frame was not a feasible solution.

(Return to timeline)

3.5 Summer 2020: MRTK overhaul

The NOMR VLE as it existed in Winter 2020 had a number of UX issues, largely around interaction with objects in the scene. It used the same toolkit (VRTK) to drive user-object interactions as Gauss’ Lab did. I ultimately realized that VRTK was too rigid and difficult to debug to be a sustainable solution, so I essentially re-created NOMR from scratch with a different toolkit.

VRTK is a community-supported toolkit, which has pros and cons. The cons outweighed the pros for my use case. At the time, Unity was in the process of rolling out preview versions of its own Unity XR Interaction Toolkit (UXRIT). Initial tests implementing NOMR's interactions with UXRIT revealed multiple severe limitations, most notably that it was prohibitively difficult to modify its source code to account for use cases not considered by its designers. I ultimately ended up switching to the Mixed Reality Toolkit (MRTK), an open-source toolkit actively developed and maintained by Microsoft's HoloLens team. At the time, a third-party patch let it be used with Oculus Quest; today, Quest is supported by MRTK out of the box.

During this time, I published the first PERC proceedings paper (Appendix C [7]) on NOMR, documenting a couple of student groups' behavior from the pilot test. I gave a couple of talks at the associated AAPT Summer Meeting and PERC. Referring to all of this work as "the VR labs Jared is making" became cumbersome, so I created the NOMR title, acronym, and cover art/logo (Figure 3.10).

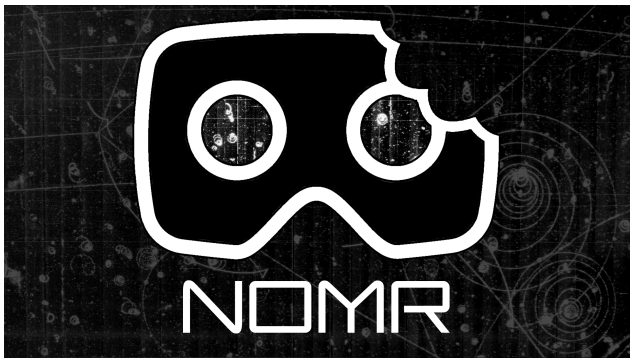


Figure 3.10: The NOMR cover art and logo are shown.

(Return to timeline)

3.6 Fall 2020: Hybrid implementation of NOMR in PHYS 231

Remote labs during COVID were, frankly, miserable. In Fall 2020, over 100 students needed to take PHYS 231: Introduction to Experimental Physics, a required course for applied physics

majors. We did our best to make these remote labs better by adapting NOMR to use for a majority of PHYS 231 lab activities.

We broke each of the seven sections into three groups of 6-8 students. Each week of Fall quarter, we had each group send a representative to lab in person, such that three students operated the lab apparatus and streamed it over Zoom to their respective lab groups. I scraped together all of the webcams, audio headsets, laptops, etc. to make this setup work, and was present for all lab sections for all but the last 1-2 weeks of the course. I set up OBS (Open Broadcasting Software) to let remote students see both the VR headset mirror and their lab partner, and remote desktop control software such that I could troubleshoot issues for students from across the room to maintain social distancing. I got precious little else done that quarter, but it was worth it to enable the one and only in-person learning experience offered by the UW Physics Department that quarter.

The NOMR labs shone as remote lab solutions. To a student in a VR headset, a collaborator sitting next to them and a collaborator a thousand miles away over Zoom are both just a voice in their ear. As their only modality of communication is audio, it makes little difference whether the students are co-located or not. From the course evaluation:

- The VR labs were pretty interesting. There was no “right” answer so messing around with that taught me a lot about the experimental process.
- VR helped me to come up with ways to test theories instead of just looking at simple, already done tests.
- I felt that the VR labs were more interesting and helpful than the hands-on labs. The VR labs allowed me to approach the problems posited without preconceived notions about what to expect, which let me take a more observational mindset towards the experiment.
- The VR labs were a really unique and cool way to get around the remote learning barrier. Good or bad, there was an element of mystery to the physics within the program that made them kind of unfamiliar to deal with when compared to hands-on labs.
- The VR labs felt more engaging (from the online side) because both online and remote people saw the same thing.
- VR Labs by far, was the best possible avenue for online labs.

(Return to timeline)

3.7 Winter 2021: Manifold Lab and remote PHYS 142

The hybrid implementation of NOMR worked fairly well in PHYS 231, so we attempted to do the same for PHYS 142, honors introductory electromagnetism, in Winter 2021. Unfortunately, so few students were able to come to campus in person with regularity that the hybrid solution seemed infeasible.

So I adapted NOMR to be used in a completely remote context. One student per lab group came to the Physics Department at the start of the quarter to borrow a VR headset.[4] I met with each "VR Operator" the week before the NOMR labs began to help them set up the screen mirroring software on their computer and make sure they could stream the headset display over Zoom. Each week, they would operate the lab "apparatus" as it were at the directions of their team members over Zoom. As needed, I had the VR Operator students install updates to the NOMR software.

During this quarter, I developed and piloted the Manifold Lab, documented in Section 2.7.3. All of the upgraded tools students could request for the last phase of the Manifold Lab were more or less developed over a weekend. Frankly, it is a miracle that they all worked as well as they did (read: not perfectly, but sufficiently). Since I was developing the tools anyway, I gave the PHYS 142 students a list of tools I was planning to implement, but left the door open for them to request custom tools not listed. This gave rise to the timed physics pause function now available in the modern Manifold Lab.

The Manifold Lab went swimmingly, from my perspective. I had been dreaming of making a NOMR activity where students collaborated between groups to iteratively create and refine models since NOMR's pilot test in March 2020. I replicated Hu and Zwickl's [25] experimental physics epistemology study with the PHYS 142 students and wrote up its results and a description of the Manifold Lab for the PERC 2021 conference proceedings [8]; see Appendix C for a copy of the proceedings paper.

---

[4] Students paid a $50 deposit that was returned when they brought the headset back at the end of the term.

(Return to timeline)

3.8 Spring - Summer 2021: Optimization, multiplayer, and tool overhaul

As well as the Manifold Lab went, it also revealed a number of areas of improvement for the NOMR VLE. Together with a team of PHYS 499 students,[5] I worked on several major improvements and additions to the NOMR VLE:

EMP Optimization: The computational expense of simulating NOMR particles' interactions scales with the square of the number of particles in the scene; in CS terms, it has time complexity $\mathcal{O}(n^2)$. This meant that, as implemented, having more than ~ 60 particles spawned at once would slow down the simulation to the point where it made for an uncomfortable VR experience. To speed things up, I rewrote EMP to use the Burst compiler, part of Unity's Data Oriented Tech Stack (DOTS). As far as I am concerned, the Burst compiler is a magic black box that speeds things up by about an order of magnitude once it is cajoled into working. This required me to rewrite EMP's field calculation functions to use the Jobs system, which makes it more difficult to implement new field functions with variables other than particle separation. It was worth the tradeoff: Now, up to $\sim 200 - 250$ particles can exist in the NOMR VLE at once before it begins to noticeably slow down. I figure if students are creating that many particles, it could only be because they are actively trying to find the limitations of the VLE, as opposed to any purpose of pedagogical significance.

Multiplayer: I implemented a Windows/Mac/Linux desktop "spectator" client that allows non-VR users to control their own camera within the VLE independent of the VR operator's view. It ultimately didn't get used because it was too unreliable on the lab computers and became less helpful once UW returned to in-person instruction in Fall 2021. I used Normcore as the multiplayer networking package.

Tool overhaul: The measuring tape was clunky to use, so I made it better. In place of hard spheres at each end of the measuring tape that could not be placed inside particles, I added flat discs

[5] Many thanks to Darcy Chen, Michael Raquel, Boris Chkodrov, Adam Tazi, Igor Banin, and James Nelson!

of a slightly larger diameter than the particles which could be placed inside particles. The tool readouts tended to overlap each other as they could not be placed independently of the tools they were associated with. I added grab-and-drop functionality to tool readouts such that each readout would begin by moving with its associated tool. If a readout is grabbed and placed elsewhere, it will remain where it is placed even when its associated tool is later moved. A pin button lets the user restore the tool-following state.

Additionally, I started to make use of Oculus' content distribution network (CDN) during this time. Previously, all builds of NOMR had to be manually side-loaded onto each headset every time a new build was created. This process was cumbersome at best. The Oculus Developer Dashboard has a Release Channels feature that lets a developer grant access to an app to users based on a list of email addresses. Once a user accepts an invitation, NOMR is downloaded through the in-headset Quest store like any other published Quest app. When a new build is created, the developer uploads it to the release channel, and the CDN pushes that updated build to all headsets with NOMR installed; the update occurs automatically so long as the headset is connected to the Internet.

(Return to timeline)

3.9 Fall 2021 - Summer 2022: NOMR settles in

UW returned from COVID-imposed distance education to in-person instruction in Fall 2021. At this point, NOMR became part of the permanent PHYS 122 lab curriculum. In every implementation of NOMR up to this point, I had been in the (often Zoom) room for every moment NOMR was in use. This changed in Fall 2021 by necessity. Prof. Kazumi Tolich was the lab instructor for PHYS 118/122, where NOMR labs were used. I was the head TA for PHYS 122 Fall 2021 - Spring 2023. In Fall 2021 and Winter 2022, Prof. Tolich and I collaborated to run the TA trainings for the NOMR labs. She took over iterating on the lab manuals accompanying NOMR, though I contributed some feedback and content each quarter. As of Spring 2022, my role in the TA preparations meetings for NOMR had dialed back to that of a senior TA. I still spoke up during the TA training fairly regularly, but my input was a value add more than a necessity.

The ease of supporting remote participants in NOMR labs continued to be particularly helpful

for Fall 2021 and Winter 2022, when it was still very common for students to be taking courses remotely or self-isolating due to a COVID exposure. Compared to hands-on apparatus-based labs, it is much easier for remote participant(s) in NOMR labs to be actively involved in designing and carrying out the experiment.

I took my general exam in November 2021.

Prof. White Brahmia and I co-wrote an NSF IUSE grant proposal over January-June 2022 to fund the further development and assessment of NOMR. The grant was rejected in January 2023.

(Return to timeline)

3.10 Summer 2022 - Present: Evaluating NOMR and handing off development

Dave Aplin and Eddie Mendoza, the UW Physics introductory lab staff, worked with me to submit a proposal to the UW Student Technology Fee committee in Spring 2022 to establish a group of paid undergraduate VR developers to continue NOMR's development. The project soft-launched in Summer 2022 with three student developers: Alnis Smidchens, Felicia Tsai, and Bryce Larson. In Fall 2022, the NOMR Student Developer Group (SDG) kicked off in earnest and has only continued to grow since. The SDG now handles the vast majority of continued development of the NOMR VLE and will continue to do so past my graduation.

Washington State University (WSU)'s introductory physics lab manager, Jacob Turner, introduced NOMR to the introductory electromagnetism lab course at WSU Pullman in their Fall 2022 semester. It remains in use at the time of writing.

I turned my focus to collecting and analyzing data to assess NOMR's effects on student learning. Over Summer 2022, Prof. White Brahmia and I developed the research plan that would become Chapter 4. The data were collected in Fall 2022, and I spent much of Winter and Spring 2023 writing up the results.

I ran four-hour workshops on NOMR at the 2023 Winter and Summer American Association of Physics Teachers meetings; they were very well-received by attendees.

In Fall 2023, PHYS 231 students who had seen NOMR labs in PHYS 122 were asked to engage in some theory development with regard to minty particles' interactions. Given their mathematical

model for the force between two minty particles, what would an expression for a minty field look like? A minty potential? We were surprised to find that these students, most of them junior and senior physics majors, could not figure out how to begin to approach these questions. It could be that these students do not understand the ideas of field and potential as mathematical constructs independent of gravitational, electric, and magnetic phenomena. I believe this to be a very interesting question to explore and highlight it as one that NOMR uniquely enables us to ask, with hope that another enterprising student will pick it up.

(Return to timeline)

Chapter 4

EVALUATION

Modeling novel physics in virtual reality labs: An affective analysis of student learning

By Jared P. Canright and Suzanne White Brahmia.

Accepted for publication in Physical Review: Physics Education Research in November 2023.

4.1 Abstract

We report on a study of the effects of laboratory activities that model fictitious laws of physics in a virtual reality environment on (1) students' epistemology about the role of experimental physics in class and in the world; (2) students' self-efficacy; and (3) the quality of student engagement with the lab activities. We create opportunities for students to practice physics as a means of creating and validating new knowledge by simulating real and fictitious physics in virtual reality (VR). This approach seeks to steer students away from a confirmation mindset in labs by eliminating any form of prior or outside models to confirm. We refer to the activities using this approach as Novel Observations in Mixed Reality (NOMR) labs. We examined NOMR's effects in 100-level and 200-level undergraduate courses. Using pre-post measurements we find that after NOMR labs, students in both populations were more expertlike in their epistemology about experimental physics and held stronger self-efficacy about their abilities to do the kinds of things experimental physicists do. Through the lens of the psychological theory of flow, we found that students engage as productively with NOMR labs as with traditional hands-on labs. This engagement persisted after the novelty of VR in the classroom wore off, suggesting that these effects are due to the pedagogical design rather than the medium of the intervention. We conclude that these NOMR labs offer an approach to physics laboratory instruction that centers the development of students' understanding of and comfort with the authentic practice of science.

4.2 Introduction

This study seeks to characterize aspects of student learning that are both highly valued [31] and challenging to assess. In the context of experimental physics courses and using a virtual reality (VR) environment, students engage in activities with novel force laws that are designed to meet a need for introductory laboratory activities that deepen undergraduate physics students' understanding of the process of generating *new* knowledge in science, and the quantitative scientific scrutiny involved. The objective of this study is to better understand the extent to which exploring novel physics, made possible through the use of immersive technologies, can render students more expert-like in their beliefs (1) about how scientific knowledge is generated and (2) in their capacity to produce scientific knowledge.

In light of the physics education research community's current understanding that laboratory instruction is not an effective means of teaching conceptual content [48, 24], we instead seek to use labs as a place where students engage in the authentic practice of science, equipping them with the tools to understand the world through an empirical lens in alignment with the AAPT's recommendations for undergraduate physics lab instruction [31]. We aim to foster an expertlike understanding of the role and process of experimentation [27, 16], build their confidence in their ability to design, perform, and interpret experiments [3, 18], and keep them actively engaged through the whole process [20].

Understanding how to provide engaging opportunities for students to develop mathematical models of novel phenomena in a teaching laboratory is a difficult open problem in physics pedagogy [31, 48, 53]. This kind of divergent, creative activity is fraught with challenges related to student autonomy and safety, opportunities for meaningful contexts [53], and the expertise of instructors to engage in a manner that responds to what is happening within each group. These issues are particularly challenging in large-enrollment courses where labs are commonly taught by inexperienced undergraduate and graduate teaching assistants (TAs).

Our framework for designing activities in which students learn to generate models is the Investigative Science Learning Environment (ISLE) approach [16, 5, 14, 15, 17, 6] by Etkina, Brookes,

and Planinsic. In ISLE, model generation happens during what the authors have named *observational experiments*, where students engage in open-minded exploration with the goal of developing a model for an unknown phenomenon. This phase of the ISLE process, itself a simplified but authentic representation of the scientific process, is followed by iteratively testing, revising, refining, and applying the model. In our approach, we alter the language slightly from Etkina to optimize transparency for the students of what they are doing. We refer to the processes of *model-generating* and *model-testing*, rather than observational and testing, experiments. Note that, in context of ISLE, the terms "model," "explanation," and "hypothesis" are interchangeable [19].

Model-generating experiments involving novel scenarios are difficult to create, especially in introductory courses. In many cases, experimental physics questions that reasonably could be investigated at the introductory level are well-known with easily-Googled answers. In presence of known answers, students tend to hold those in highest regard, seeking to confirm known answers above anything and everything else, even in the face of contradictory data [52] or explicit instruction to the contrary. Thus, any access to a "right answer" can derail efforts to engage students in authentic model generation. We address this expectation of getting a right answer by putting students into a different universe with *new* physics that builds on Newton's laws and fundamental conservation laws – where neither they, nor their textbook, nor Google, nor their TAs have a ready-made model at hand. The problem of shifting students' mindset toward generating new models becomes trivial when there are no existing models to confirm.

The activities described in this paper have been part of the University of Washington (UW) introductory physics curriculum for over two years. In the Novel Observations in Mixed Reality (NOMR) labs [7, 8], students explore real and fictitious physical phenomena in an immersive 3D environment. Instructors are struck by the ways that students mature as scientists through these labs, an impression that is not easily quantified. The following is an excerpt from a post-course survey that is fairly typical at the sophomore level, and representative of those most impacted at the introductory level.

"VR labs were fantastic for learning how to effectively approach a physics situation

> where I didn't already know what would happen. In most experiments I have done in previous courses, I had learned what to expect before I was actually making observations and collecting data, so this course helped me learn a new way to approach experiments."

This study is a step toward characterizing this kind of intellectual growth we observe in many students. The work is situated in efforts across the physics education community to find and adapt affective assessment tools beyond standard course evaluations. Our study seeks to establish whether students' belief in their ability to do physics and their sense of belonging in physics grow along with their understanding of the nature and role of experimental physics to generate new knowledge. We assess the impact of the intervention on students': epistemology about experimental physics; physics self-efficacy; and engagement in the learning process. This study contributes to the ongoing research into assessment of what students take away from effective laboratory instruction [55, 54]. Specifically, we focus on the following research questions:

RQ1 What changes are observed in students' epistemology about experimental physics as a result of the NOMR labs?

RQ2 What changes are observed in students' physics self-efficacy in experimental physics as a result of the NOMR labs?

RQ3 To what extent are students productively engaged in the NOMR activities, and how does that engagement compare with the hands-on labs in the same course?

4.3 Background

4.3.1 ISLE

We begin by considering what lab activities reflecting the real-world practice of science look like. The ISLE approach to physics education [16, 5, 14, 15, 17, 6] prioritizes epistemologically authentic investigation of physics as a means to develop students' scientific abilities [18] and habits of mind. Teaching students to think like expert physicists takes priority over covering conceptual content.

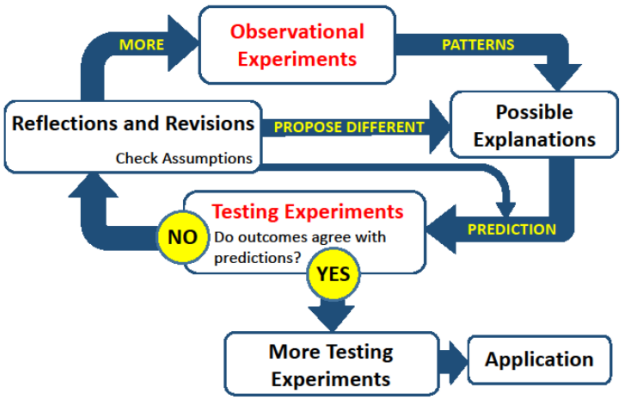


Figure 4.1: The ISLE process is a simplified representation of the real-world practice of science, iteratively generating, testing, rejecting, and refining models to empirically create and validate knowledge.

Three types of experiments form the core of ISLE instructional activities, related to each other by the ISLE process (Figure 4.1):

Model-generating experiments Labelled in the diagram as Observational Experiments, students engage in open-minded exploration of a previously unknown physical phenomenon. They make note of patterns in the phenomenon's behavior and devise explanations for those patterns. These patterns become mathematical models and mechanistic explanations of the phenomenon. The models created in model-generating experiments form the basis of students' knowledge of the phenomenon.

Testing experiments A model from a prior model-generating experiment is tested. Students design an experiment with well-determined independent, dependent, and controlled variables to test a prediction about the outcome of an experiment that follows from the model in question. They run the experiment, collect and analyze their data, and judge whether the outcome is consistent with the prediction. If so, they have supported, or failed to reject, the model. If not, the model is rejected.

Application experiments Students apply tested models to determine the value of unknown phys-

ical quantities or solve practical problems.

In the ISLE approach, content and process are considered to be inextricably paired; these three types of activities are how students uncover new physics content. Students encounter physical phenomena for the first time through hands-on experimentation, and only after identifying patterns and developing their own explanations do they read about the phenomena in their text.

ISLE-approach labs consist primarily of questions to guide students' thoughts rather than dictate them. In this way, students learn to ask and answer questions in the way that a scientist would. This process is guided and refined by scientific abilities rubrics [18] used to assess and give feedback on their work.

4.3.2 Affective measures

This study employs three different research-validated surveys to explore different aspects of students' engagement and learning. Table 4.1 gives the name, abbreviation, target metrics, format, and administration schedule for each survey.

Survey		**Target Metrics**	**Format**	**Schedule**
Lab Epistemology Survey	LES	Student attitudes and beliefs about the nature and role of experimentation in physics	Five open-ended short-answer questions	Pre-survey at the beginning of the term; 100-level post-survey after NOMR activities concluded; 200-level post-survey after course final
Physics Identity Survey	PhIS	Degree of students' self-identification as a physicist, interest in physics, and belief in their ability to practice and succeed at physics	Six 6-point Likert items (self-perception and interest); Five 7-point Likert items (self-efficacy)	
Flow Survey	FS	Degree and nature of students' engagement with the week's lab activity	Seven 7-point Likert items	Weekly at the conclusion of each lab activity

Table 4.1: The characteristics and administration schedule of each survey used in this study are summarized.

Epistemology about experimental physics

The Lab Epistemology Survey (LES) originally developed by Hu and Zwickl [25, 26] is used in this study as a measure of students' epistemology about the role of experimental physics in class and in the world. The LES was developed as an instrument to characterize the beliefs held by physics students at introductory undergraduate, upper-division undergraduate, and graduate levels about about experimentation, models, and their roles in the scientific process. This study uses the LES pre-post to assess changes in students' epistemology about experimental physics before and after students complete the NOMR labs. Measuring changes in students' epistemology gives us a window into whether they are adopting the beliefs, attitudes, and mindset about experimental physics characteristic of expert physicists.

The LES is composed of six open-ended questions, accompanied by a codebook used to identify themes in student responses in a consistent and reproducible way. We focus on the first two questions:

LES1 In your opinion, why are experiments a common part of physics classes? Provide examples or any evidence to support your answer.

LES2 In your opinion, why do scientists do experiments for their research? Provide examples or any evidence to support your answer.

Novice responses to these LES items exhibit an almost singular focus on the idea that experiments in instructional labs exist to supplement conceptual learning or test theories (using a layperson's understanding of the term "theory"). The term "theory" is somewhat vague here: It has a specific definition in the context of physics, but is used in a lay sense by students, often translating to "anything that is not an experiment" or "what we know from the textbook or lab manual." Expertlike responses more frequently acknowledge the role of in-class experimentation in the development of scientific abilities and as a means to better understand the scientific process. With regard to experiments in scientific research, novices tend to focus on the notion that experiments exist to test theories. Experts more frequently cite the creation of new models and the iterative nature of experimental model development as purposes of experiments in research.

Item	Code	Definition	Example Student Response
LES1: Why are experiments a common part of physics classes?	*Modeling*	Experiments in class let students develop their own models for phenomena, discover things on their own, and/or develop their own ideas.	Because it helps show the process of developing a model, rather than just taking it as fact and using it to solve problems. By studying "mystery particles" in lab, we had to experiment and develop our own observations.
	Scientific abilities	Experiments help cultivate students' scientific abilities, such as experimental design, data collection, and data analysis skills.	Experiments provide a way to provide reasoning skills as applied to physics, of which experiments *[sic]* reasoning is needed to have problem [solving] skills
	Model testing	The purpose of doing physics experiments is to prove, support, or test a model.	Experiments are necessary to test theories. Theories cannot be made into laws without testing.
	Supplemental learning	Experiments provide supplemental learning experiences for concepts and theories.	Experiments are a common part of physics courses because they help you understand the concepts we are learning.
LES2: Why do scientists do experiments for their research?	*Discoverment*	Experiments contribute to some aspect of the iterative and generative nature of the scientific process aside from testing an existing model.	Scientists in the real world are consistently working to provide new findings that deepen our understanding of the world.
	Model testing	The purpose of doing physics experiments is to prove, support, or test a model.	You cannot confirm a hypothesis without performing experiments.

Table 4.2: LES items, the codes associated with each, and an example response tagged with each code are shown. The example responses for each code were selected such that each example was assigned only the associated code. Many responses in the data were assigned more than one code. *Scientific abilities* and *Supplemental learning* are drawn from Hu and Zwickl's original codebook [25]. *Model testing* is equivalent to the original codebook's *Theory testing*. We added the *Modeling* code in response to the recurring presence of its ideas in our dataset and its relevance to our research questions. The original codebook's *Discovery* and *Theory development* codes were merged to create *Discoverment*.

The original LES [25] included the codes *Theory testing* ("The purpose of doing physics experiments is to prove a theory or test a hypothesis.") and *Theory development* ("Experiments inspire the development or improvement of theories."). Due to the vague nature of the term "theory," it is unclear how Hu and Zwickl drew a distinction between "theory" and "hypothesis" as used by students. To use the term "theory" in our analysis, we would need to establish our own definition, at risk of misrepresenting students' responses in a replication study.

In alignment with ISLE, we remove references to the term "theory" in favor of the term "model." A model is a foundational concept in ISLE, used heavily throughout both populations' lab activities. The modified codes we use in our analysis are ***Model*** *testing* ("The purpose of doing physics experiments is to prove, support, or test a model.") and ***Model*** *development* ("Experiments inspire the development or improvement of models.").

The distinction between *Model development* and *Discovery* (Original definition: "Experiments help investigate unknowns.") is subtle and not one we are interested in probing; rather, we are more interested in understanding whether students' responses reflect any acknowledgement at all of steps of the ISLE process aside from model testing. Instead, we collapse the two codes into a single *Discoverment* code: "Experiments contribute to some aspect of the iterative and generative nature of the scientific process aside from testing an existing model."

The final code list is given with definitions and examples in Table 4.2.

Physics self-efficacy

In context of the physics classroom, self-efficacy refers to students' belief in their ability to practice and succeed at physics. Developing students' self-efficacy is a primary goal for our laboratory instruction, as we want students to walk away from the course with confidence in their ability to design, perform, and interpret experiments [3, 18].

This study employs the Physics Identity Survey (PhIS), which we adapted from a science identity survey administered to middle school biology students to evaluate shifts arising from their participation in an immersive virtual lab [42]. The original survey was developed through the lens of Hazari's science identity framework [22].

The PhIS is divided into two sets of items probing (1) self-efficacy, and (2) physics identity and interest. We focus on the self-efficacy items, listed below, each on the scale [1: Not at all Confident – 7: Completely Confident]:

PhIS1 How confident are you that you can design an experiment to answer a scientific question in physics?

PhIS2 How confident are you that you can look at data that you collect and characterize its patterns mathematically?

PhIS3 How confident are you that you can understand the kinds of problems that experimental physicists would investigate?

PhIS4 How confident are you that you could contribute to a team of physicists investigating an experimental physics problem?

PhIS5 How confident are you that you can defend your data analysis to a team of expert physicists?

These items were adapted by swapping out learning goals of the ecosystems biology course of the original study for learning goals of the lab courses of concern in this study, e.g., designing an experiment to answer a scientific question in physics and mathematically characterizing patterns observed in data.

We validated the Likert scale items comprising the PhIS in accordance with Adams and Weiman's recommendations for the development of formative assessment instruments [1]. We conducted think-aloud interviews with eight 100-level physics students. Participants were asked to rate each Likert-scale item and explain their choice as they did so. Participants were recruited through an announcement over the course's web page and incentivized incentivized to participate with $20 gift cards. The interviews were conducted online, audio recorded, and transcribed by Otter.ai and subsequently hand-corrected.

The interview transcripts were examined to assess the alignment of students' reasoning for the responses they chose with the construct each item was meant to assess. Students' understanding of each item reflected our expectations, and their reasoning for each choice revealed nothing unexpected. The validation interview results did not lead to any modification of the PhIS.

Flow as a measure of engagement

We use the psychological theory of flow pioneered by Csíkszentmihályi [11] as a lens through which to examine students' engagement with class activities. Known colloquially as being "in the zone," flow is described as a state in which one is completely absorbed in an activity for its own sake, where one action leads smoothly into the next, and one's sense of time becomes distorted. A balance between the person's self-perceived skillfulness at and the challenge posed by an activity is instrumental to achieving a flow state; great challenge must be met with commensurate belief in one's own skill. It is in this state that the most effective learning happens [12].

Csíkszentmihályi identified seven conditions for a person to achieve flow:

1. They know what to do (a clear goal).
2. They know how to do it.
3. They are receiving clear and immediate feedback to know how well they are doing.
4. They know where to go (if navigation is involved).
5. They see what they are doing as challenging.
6. They are confident in their ability to complete the task.
7. Their environment is free of distractions.

As the flow model is fundamentally one of engagement with an activity, it has utility as a measurement of students' engagement with the learning process [12, 41, 29]. Active engagement is key to learning [20]; conversely, even well-designed instructional activities with epistemologically authentic inquiry as in ISLE cannot reach students who are not engaged in the learning process.

A subset of the flow conditions comprise an effective basis for maintaining task involvement: A learner needs feedback, confidence in their ability to complete the task, and an environment free of mental distractions. To stick with a task to completion, it is critical a the student's self-efficacy is great enough that they believe they can do so [3]. To support this belief, they need clear feedback to know how well they are doing and what the next steps are.

Rebello and Zollman note [41] that the zone of proximal development [58], the optimal adaptability corridor [46], and flow are all representations of a balance between a learner's skill and

challenge. To express the optimal adaptability corridor's dimensions in terms of flow, horizontal transfer (efficiency) maps to skill, and vertical transfer (innovation) maps to challenge. Flow comes in as a means to tie this balance to other affective elements of the student experience, unifying a number of affective constructs in educational psychology under one quantifiable umbrella.

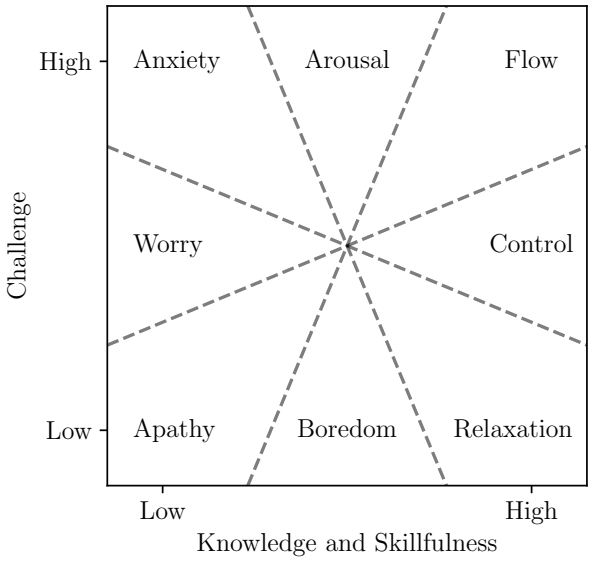


Figure 4.2: The eight-channel model of flow.

Massimini and Carli's efforts [35, 34] to develop a quantitative instrument to measure flow led to the eight-channel flow model we use in this study. One begins by constructing a mental state diagram, hereafter referred to as a flow plot, with perceived knowledge and skillfulness on the horizontal axis, and perceived challenge on the vertical axis. Each flow plot is divided into eight channels (Figure 4.2), representing different relative combinations of challenge and skill. The top-right channel, Flow, is the most productive, representing a great challenge met with commensurate skill. Flow's neighbor channels Control and Arousal represent less challenge and less skill than Flow, respectively. Flow, Control, and Arousal are considered productive channels for learning [12]. The Relaxation channel represents a surplus of skill and dearth of challenge; its mirror

channel Anxiety represents an extreme challenge one feels poorly equipped to handle. The least productive states of Worry, Apathy, and Boredom fill out the lower regions of the challenge-skill space, with skill and challenge both insufficient to support productive engagement.

Karelina, Etkina, Bohacek, Vonk, Kagan, Warren, and Brookes [29] used the eight-channel model to compare students' engagement with content-equivalent ISLE-aligned hands-on and video-based labs; we follow much of their methodology in our study of students' engagement with VR labs.

The Flow Survey (FS) is drawn from Karelina *et al.*'s adaptation of a subset of items from the psychometrically-validated Flow State Scale [28]. The FS uses items from their adaptation with minor wording changes for our experimental context.

It consists of seven 7-point Likert scale items:

F1 To what extent was the instructor's assistance needed? [1: Not at all – 7: A lot]

F2 To what extent did you know what to do (goal of the task)? [1: Not at all – 7: A lot]

F3 To what extent did you know how to do it? [1: No idea – 7: Completely]

F4 To what extent did you know how well you were doing? [1: No idea – 7: Completely]

F5 To what extent was the lab challenging? [1: Not at all – 7: Extremely]

F6 To what extent did you feel knowledgeable and skillful during the lab? [1: Not at all – 7: Extremely]

F7 To what extent was the lab fun and interesting? [1: Not at all – 7: Extremely]

We follow Karelina *et al.*'s analysis methods to create flow plots using two items: **F5** as a measure of percieved knowledge and skill on the horizontal axis, and **F6** as a measure of challenge on the vertical axis. Students who give a high score to both items are understood to be in a flow state.

Karelina *et al.*'s study compared average responses along each axis of the flow plot with two-tailed paired t-tests to determine differences in students' perceived skill and challenge between video and hands-on treatment groups. These quantitative comparisons were backed by visual comparisons of which channels students tended to fall into in each treatment group.

We also make use of **F7** ("...fun and interesting?") to characterize the effect of the novelty

of VR on students' experience. Weekly measures of this item allow for comparison across lab activities, e.g., comparing hands-on labs to VR labs.

4.4 Methods

4.4.1 Structure of NOMR labs

The intervention examined in this study uses VR labs to allow students to experience and analyze physical laws in the context of particle interactions that do not exist in nature or on the Internet. We include the constraint that they be consistent with our universe's physics so that students can rely on their extant physics knowledge when reasoning in the VR space. These fictitious physical laws can be construed as hypothetical mathematical variations on Coulomb's Law. A selection of fictitious phenomena are described in more detail in [7].

The virtual apparatus is designed such that it does not give perfect answers; experimental uncertainty is still very much present even in the simulation. The "right answers" programmed into the simulation are never shared with students or their TAs, such that the only "right" answer is the one that students can make the best case for.

In the virtual lab space, students can access force and distance measurement tools and a supply of particles (modeled as hard spheres) exhibiting the behavior(s) they are investigating. These particles can be moved around the space freely, as well as be fixed in place individually. To facilitate creating static arrangements of particles, physics can be temporarily paused in the entire space. As there is no copy of the lab manual nor any means by which to record data while in the headset, the operator relies on their group's interaction and record-keeping. Each group of 3-4 students shares one VR headset, with its display mirrored onto a lab computer.

Multiple instructional practices are in place to combat gendered task division common to inquiry-based physics labs [13, 23]. All groups complete a teamwork agreement at the beginning of the term outlining expectations and norms for their interactions in and out of class. The NOMR lab manuals prompt students to take turns using the headset at multiple junctures. Students are encouraged to have each member of the group use the headset to collect data, as a means of

obtaining multiple measures from which to determine a central value and uncertainty for each of their measurements.

The NOMR labs described in this study are used in two instructional contexts: in introductory calculus-based physics, and in a sophomore-level lab for applied physics majors. The first lab encountered, called Charge and Mint, is comprised of two activities. Introductory students complete these components as two separate labs (VR1 and VR2; see full lab titles and schedule in Table 4.3), and students in the advanced course complete both components in a single lab session (VR1+2).

First, groups design and conduct an experiment to test whether virtual analogues of electrically charged particles follow some re-scaled version of Coulomb's law. This lab serves to familiarize students with the VR environment and doubles as an opportunity to teach (or review) data linearization.

Second, they take qualitative and quantitative data to create an empirical model for the interaction between fictitious *minty* particles, which behave according to an unknown force law. That force law is not included here in an effort to keep it out of print; instead, we present a handful of observations students might make, and leave the specifics of the model to the reader's imagination:

- Minty particles repel when they are near each other, and attract when they are far away. A turnaround point where the force is zero exists at a certain separation between particles.
- If two minty particles are brought as close as possible to one another and released from rest, they appear to undergo oscillatory motion. Considering the full range of distances achieved in this motion, the range of distances for which the force between the particles is repulsive seems to be shorter than the range of distances for which it is attractive.
- Beyond the repulsive region, no matter how far apart two minty particles are moved, they continue to exert upon each other a substantial attractive force which increases with distance.

Charge and Mint is used as a preparatory lab to get students familiar with the virtual learning environment and comfortable with the idea of developing a mathematical model for a completely unknown phenomenon. Once the preparatory lab is complete, students are given a new, more complex phenomenon to explore and model, without any phenomenon-specific scaffolding. This final lab takes two forms: the one-week Exotic Matter Lab for introductory students (lab VR3),

and the three-week Manifold Lab (VR3-VR5) for advanced students.

The Exotic Matter Lab's content is identical to the first week of the Manifold Lab: Every group is assigned a different set of fictitious phenomena (referred to as a *scenario*). This phase allows for differentiated instruction as the TA assigns scenarios at the start of class based on their impressions of each group's strengths and weaknesses and the nature of the challenge each scenario poses. Each scenario contains up to three distinct types of particles, all visually identical on creation, picked from at random whenever the user creates a new particle. The phenomena underpinning each scenario, and the subject of students' inquiry, are the force laws governing the interactions between the particles. In most cases, a single force law dictates the interaction between each pair of particles, though students may develop different valid interpretations supported by their data. The force laws programmed into NOMR are never shared with students or TAs.

In this model-generating experiment, students are told that they have at most three distinct types of particles in their scenario and given tools to label particles and temporarily remove particles from play. Their goals are to determine how many types of particles they have, develop a procedure for identifying an unknown particle, and come up with a testable empirical model describing some subset of the behaviors they observe. Students write up their findings in a full lab report. For introductory students, this model-generating experiment report marks the end of their foray into VR.

Advanced students working through the Manifold Lab instead submit reports describing their model-generating experiments and resulting models to a class-wide repository. During class in the second week, each group selects another group's report describing a model of a scenario they had not yet interacted with themselves. They write and submit a proposal before the third week of lab, describing an experiment to test the other group's model. These experiments are carried out in the third week, and their results presented in an oral talk symposium in the fourth week.

The Manifold Lab is presented to students with a game-like narrative in which they function as research scientists. They explore different "pocket" universes with novel forms of matter obeying fundamental force laws unknown to our universe. The privilege to conduct the second experiment with better equipment depends on applying (non-competitively) for grant funds: Before performing

the second experiment, students write a single-page "grant proposal" in which they summarize another group's findings, propose an experiment to test their model, and request additional or improved equipment within the VR lab. The instructor serves as an entity equivalent to the NSF and its reviewers: They review students' grant proposals for feasibility and work with each group to revise proposed experiments such that it is likely that each will produce a clear outcome that builds on the prior group's findings. Each group receives a few credits to spend on equipment, e.g. more precise measurement tools, a larger workspace, tools that snap to more convenient configurations, or the ability to automatically pause physics after a set amount of time. Occasionally, a clever idea from a group of students inspires the development of a new tool in NOMR, which is added to the upgrade options going forward.

This design seeks to emulate the experience of working within a professional scientific collaboration: The class as a whole collaborates by sharing data and designing experiments to test and revise each other's models. In doing so, students complete an entire cycle of the ISLE process: One group creates a model through a model-generating experiment, another tests it with a testing experiment, and those results serve to reject, revise, or further substantiate the model.

4.4.2 Instructional Context

The study activities took place at the University of Washington in Seattle (UW), a large R1 public research university in the Pacific Northwest. Of the population of students enrolled in the courses examined in this study, 65% identify as male, and 35% as female (non-binary gender identities are not reflected in UW records, and we did not solicit this information from students separately). White (41%) and Asian (36%) students make up the majority of the population, followed by students who identify with two or more races (8.7%), Hispanic or Latino,a,e students (7.6%), and Black students (2.8%).

Our data come from two physics courses at UW during Fall 2022. All instruction was held in person except in the event a student couldn't attend a lab due to illness, in which case their lab partners brought them in via video call, when possible. In both courses, groups of 3-4 students worked together for the entire quarter. Each class's lab curriculum schedule is shown in Table 4.3.

Wk	100-level			200-level		
	Lab	**Name**	**Type**	**Lab**	**Name**	**Type**
1	0	Coulomb's Law	T	E1	Electron Beam pt 1	T
2	VR1	Charge	T	E1	Electron Beam pt 2	C
3	VR2	Minty	G	N1	Nuclear Decay pt 1	T
4	VR3	Exotic Matter	G	N2	Nuclear Decay pt 2	C
5	B1	Bulbs pt 1	G	VR1+2	Charge + Minty	T / G
6	B2	Bulbs pt 2	T	VR3	Exotic Matter	G
7	B3	Capacitors	G	VR4	Manifold Proposal	C
8	(No lab)					
9	B4	Unknown Resistor	A	VR5	Manifold Testing	T
10	(No lab)			VR6	Final Presentations	C

T = Testing, G = Generating, A = Application, C = Communication

Table 4.3: This timeline of events shows the curriculum 100-level and 200-level students completed over the course of the quarter, and when each survey was administered to each population. All LES & PhIS pre- and post-surveys were administered outside of class time except for the 200-level post-test, which students completed in class after their final presentations. The labels **G**, **T**, **A**, and **C** represent model **generation** experiments, model **testing** experiments, **application** experiments, and **communication**, respectively.

100-level population NOMR was implemented during calculus-based electromagnetism, the second of a three-quarter introductory physics sequence. 467 students enrolled in Fall 2022, and 380 students consented to participate in the study. We refer to the consenting students as the *100-level population* hereafter. This course consisted largely of engineering (74%) and science (17%) students filling prerequisites for their major. Students met weekly for three 1-hour lecture sections, a 1-hour tutorial section, and a 2-hour lab section.

200-level population Introduction to Experimental Physics used NOMR as well. This course enrolled 38 students, mostly applied physics majors (55%) who typically intend to follow an industry-oriented path after graduation, alongside other physics and astronomy majors (21%), pre-science majors (13%), computer science majors with physics minors (8%), and one math major. All 38 students consented to participate in the study; we refer to them as the *200-level population* hereafter. Students met weekly for a 1.5-hour lecture section and 3-hour lab section. Eight members of the 200-level population had previously seen NOMR labs in the modern 100-level labs; all other 200-level students had taken traditional or online (due to COVID) 100-level labs, without VR.

Lab activities

All lab activities in the 100- and 200-level courses are designed in alignment with the ISLE approach. We say "in alignment" because a full implementation of ISLE requires integration of the ISLE process across all components of a course (lecture, lab, etc.), which is not the case at UW.

Every lab activity can be categorized as a hypothesis generating, hypothesis testing, or application experiment (as in Table 4.3), excluding weeks in the 200-level course dedicated specifically to writing and communication. Students' work is guided and assessed with the ISLE scientific abilities rubrics [18].

The first four weeks of the 100-level labs (labs 0, VR1-VR3) focus on particle interactions. Lab 0 is a qualitative testing experiment on Coulomb's law. Students test the effects of charge and separation on the electric force between a copper sphere and a Teflon rod. This is the simplest lab of the quarter, and deliberately so. It is the first lab of the quarter, when students are still joining the

course and switching between sections. The following three weeks (labs VR1-VR3) are NOMR labs, as described in section 4.4.1.

The remaining four weeks of labs are traditional hands-on labs exploring circuits:

B1 Students begin their exploration of circuits with a model-generating experiment seeking to develop a model to describe the behavior of battery-bulb circuits in series, parallel, and mixed configurations.

B2 Models generated in the prior week are tested against a mixed-configuration circuit. Students create predictions for the current and voltage through each element of the circuit based on their model from the prior week, build the circuit, collect data, and compare the results to their predictions. Where there is disagreement, students revisit and revise their model.

B3 Capacitors are introduced. Students perform a model-generating experiment to develop a mathematical model for the voltage across a charging capacitor.

B4 Students are given a resistor of unknown value and a model of voltage across a discharging capacitor. Using this model, students perform an application experiment to determine the value of the resistor, with uncertainty, by manipulating the model to give the resistance in terms of the slope of a linearized plot and the capacitance of the capacitor.

The 200-level course opens with the electron beam lab (E1-E2). Each group is given an electron beam apparatus (commonly called an "e/m apparatus") that fires electrons across a user-specified voltage into a helium-filled bulb subject to an approximately uniform magnetic field generated by Helmholtz coils outside the bulb. Students are asked to devise and answer a scientific question with the apparatus. Most often, this ends up being a testing experiment based on students' knowledge of electrons' motion in a magnetic field. Occasionally, it turns into a model-generating experiment if a group does not recall this model.

The subsequent nuclear decay lab (N1-N2) works in similar fashion: Students are given an apparatus, instructed in its operation, and are set loose to devise and answer a scientific question of their choosing. In this case, the apparatus is a radioactive Cs source, a Geiger-Muller tube with event counting hardware and software, a box of barriers of various material and thickness, and a stand for all of the above with slots in which to place the source and barriers.

The rest of the 200-level labs are NOMR labs documented in section 4.4.1: Charge and Mint (VR1+2) and the Manifold Lab (VR3-VR6).

4.4.3 Data Collection

Lab Epistemology Survey and Physics Identity Survey

The LES and PhIS were administered as part of the same survey in all cases.

100-level students completed the pre-survey in Week 2, after Lab 0, which was a traditional hands-on lab, and before VR1, the first NOMR lab. The survey was included as part of a timed quiz, and we recognize the time constraint may have influenced students' responses. The post-survey was included as part of an untimed reflection on the performance of their group. It was administered in Week 4 after the conclusion of VR3, the final NOMR lab for 100-level students. Therefore, the pre-post shifts reported here reflect changes in 100-level students' responses before and after only the NOMR labs.

200-level students completed the pre-survey in Week 1 before the start of classes, and completed the post-survey in Week 10, after their final presentations. The pre-post shifts reported here reflect changes in students' responses before and after all 9 weeks of labs, including non-NOMR activities.

Flow Survey

The FS was administered every week at the end of lab to both populations. We offered a small amount of extra credit for each week the survey was completed, and emphasized that the surveys would help improve the lab curriculum in future terms. In the 100-level population, 42 students completed all eight surveys; in the 200-level population, 14 students completed all eight surveys. The findings reflect responses only from these students who completed all eight surveys.

4.5 Findings

4.5.1 RQ1: What changes are observed in students' epistemology about experimental physics as a result of the NOMR labs?

Responses to **LES1** and **LES2** were coded independently by the first author and another researcher. The researchers met to reconcile disagreements after the first coding pass, with a disagreement rate of roughly 10% for **LES1** and 30% for **LES2**. After further expanding on the existing code definitions, adding examples, and adjusting the codes as described in Section 4.3.2, we reached $> 95\%$ inter-rater agreement across all codes.

The LES findings are presented as code frequencies: the fraction of responses in a population that were assigned each code. Figures 4.3 and 4.4 depict the shift from pre to post for each population-item. Each student response could be assigned no code, one code, or more than one of the codes associated with the item. As most responses were assigned one or more codes, the total number of codes is greater than the number of responses.

Cohen's d is presented for each shift, calculated from pre and post code frequencies p_{pre} and p_{post}:

$$d = \frac{|p_{post} - p_{pre}|}{\sqrt{\frac{\sigma_{pre}^2 + \sigma_{post}^2}{2}}}; \quad \sigma_i = \sqrt{p_i(1-p_i)} \tag{4.1}$$

Whether a code is or is not assigned to a given response is a binary variable, so the binary standard deviation is used.

LES1: Why are experiments a common part of physics classes?

Responses to **LES1** were assigned up to four codes, defined with examples in Table 4.2 and reproduced below:

Modeling Experiments in class let students develop their own models for phenomena, discover things on their own, and/or develop their own ideas.

Scientific abilities Experiments help cultivate students' scientific abilities, such as experimental design, data collection, and data analysis skills.

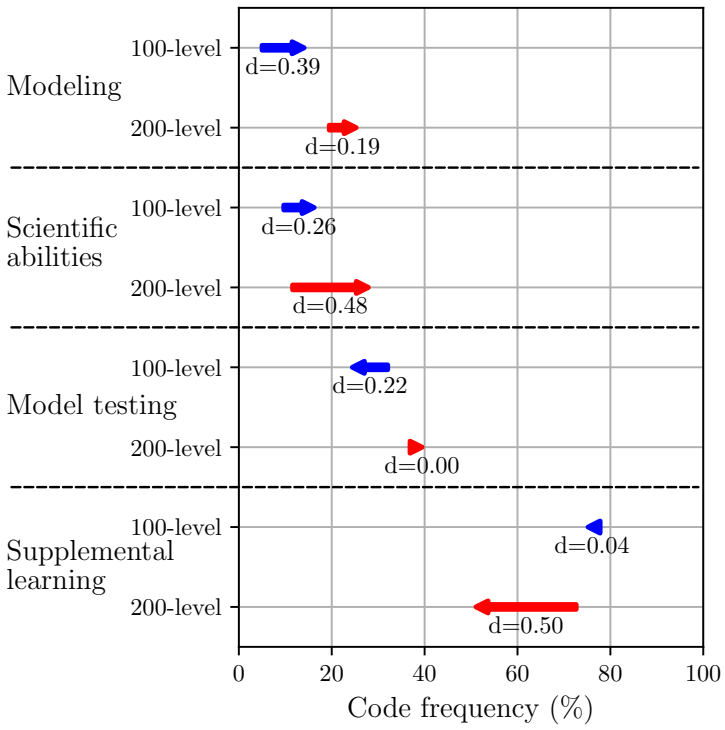


Figure 4.3: Code frequencies for both populations' responses to **LES1**. A code frequency represents the percentage of responses in a population that were assigned that code. Each response could be assigned zero, one, or multiple codes, so percentages do not add to 100%. Data from our 100-level population are in blue, and the 200-level population in red; the tail and head of each arrow represent pre-intervention and post-intervention data, respectively. This graphic represents $N_{100} = 278$ and $N_{200} = 38$ matched pre-post responses from 100-level and 200-level students, respectively. Cohen's d is calculated according to Eqn. 4.1.

Model testing The purpose of doing physics experiments is to prove, support, or test a model.

Supplemental learning Experiments provide supplemental learning experiences for concepts and theories.

We added the *Modeling* code in response to the recurring presence of its ideas in our dataset and its relevance to our research questions. It captures the creation of new models that results from model-generating experiments in the ISLE process, while *Model testing* represents the subsequent model-testing function of a testing experiment.

Example responses that were assigned each code are given in Table 4.2. A sample response to **LES1** that was assigned multiple codes follows:

> "Experiments allow us to challenge what we know while apply what we have learned. Its a new way of learning things–a more hands-on approach. We can learn about the scientific models and how experiments are designed to either explain a new phenomena or test a pre-existing model."

This response is assigned *Modeling* for the phrase "We can learn…how experiments are designed to either explain a new phenomena…" and *Model Testing* for the last part of that sentence: "…or test a pre-existing model." *Supplemental learning* is present in a couple of places: "We can learn about the scientific models…" and "Its [sic] a new way of learning things–a more hands-on approach." Finally, the clause "…how experiments are designed…" merits the *Scientific abilities* code.

Both populations' code frequencies are plotted in Figure 4.3. Comparing pre-post results, we find that the students in our study became more likely to indicate a belief that labs are meant to develop their *Scientific abilities* and give them opportunities to develop their own models (*Modeling*). The 100-level population became less likely to cite *Model testing* as a purpose of in-class experiments, while the 200-level population became dramatically less likely to cite *Supplemental learning* for the same.

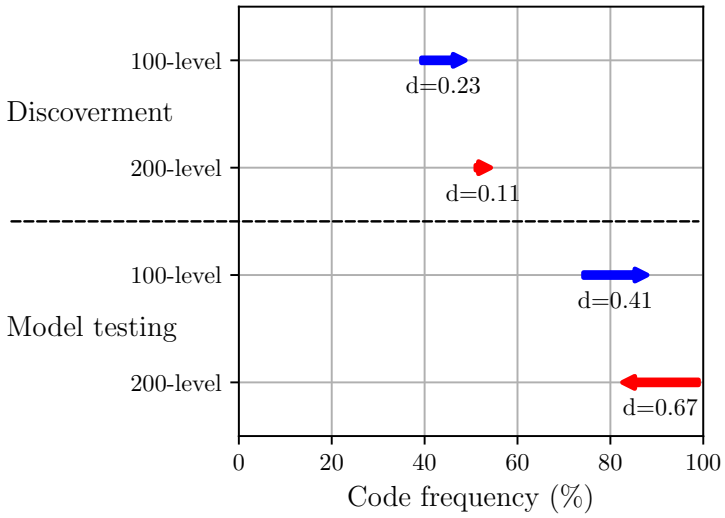


Figure 4.4: Code frequencies for both populations' responses to **LES2**. A code frequency represents the percentage of responses in a population that were assigned that code. Each response could be assigned zero, one, or multiple codes, so percentages do not add to 100%. Data from our 100-level population are in blue, and the 200-level population in red; the tail and head of each arrow represent pre-intervention and post-intervention data, respectively. This graphic represents $N_{100} = 278$ and $N_{200} = 38$ matched pre-post responses from 100-level and 200-level students, respectively. Cohen's d is calculated according to Eqn. 4.1.

LES2: Why do scientists do experiments in professional research?

Responses to **LES2** were tagged with up to two codes, defined with examples in Table 4.2 and reproduced below:

Discoverment Experiments contribute to some aspect of the iterative and generative nature of the scientific process aside from testing an existing model.

Model testing The purpose of doing physics experiments is to prove, support, or test a model.

LES2's code frequencies are plotted in Figure 4.4. Both populations acknowledged the iterative and generative elements of the scientific process (*Discoverment*) more frequently after NOMR labs than at the beginning of the course. The 100-level population's *Discoverment* code frequency increased by a significant degree; the 200-level population's frequency started higher and saw a smaller increase. The 200-level change is small enough to be statistically insignificant.

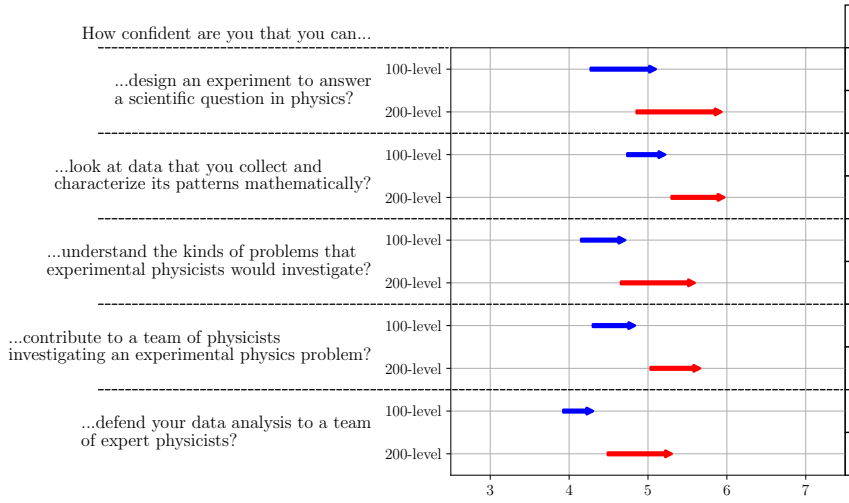
The 100-level population was assigned *Model testing* codes more frequently after NOMR labs than at the beginning of the course, jumping from 73% to 89%. The 200-level population's code frequencies saw a significant decrease in *Model testing*, moving from 100% to 82% of matched responses.

4.5.2 *RQ2: What changes are observed in students' physics self-efficacy in experimental physics as a result of the NOMR labs?*

Statistical Analysis

We assess pre-post shifts in students' responses to the Likert items comprising the PhIS by calculating p-values using the Wilcoxon signed-rank test [57]. We chose the Wilcoxon test because it is a nonparametric test and thus makes no assumptions about whether the dataset is normally distributed. We opted against using the Mann-Whitney U test as we are interested in testing for differences within paired samples. The Wilcoxon signed-rank test is a more appropriate choice than the Mann-Whitney U test, since the Mann-Whitney test is used for independent samples.

We report the common-language effect size f computed from the Wilcoxon test statistic T and the total rank sum S by the expression: $f = \frac{1+T/S}{2}$. In general terms, f tells us what fraction



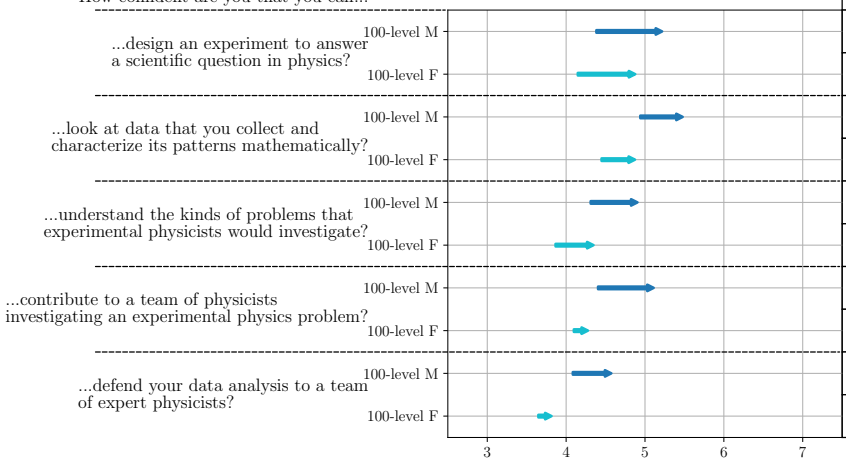
	Shift	Effect size f	p-value	T_+
...design an experiment to answer a scientific question in physics? — 100-level	0.890	0.852	2.71e-13	8057.5
...design an experiment to answer a scientific question in physics? — 200-level	1.141	1.000	5.93e-06	351.0
...look at data that you collect and characterize its patterns mathematically? — 100-level	0.545	0.721	1.31e-06	8268.5
...look at data that you collect and characterize its patterns mathematically? — 200-level	0.733	0.925	0.000184	234.0
...understand the kinds of problems that experimental physicists would investigate? — 100-level	0.622	0.770	6.02e-09	8154.5
...understand the kinds of problems that experimental physicists would investigate? — 200-level	1.000	0.930	0.000157	279.0
...contribute to a team of physicists investigating an experimental physics problem? — 100-level	0.599	0.699	6.72e-06	9122.0
...contribute to a team of physicists investigating an experimental physics problem? — 200-level	0.692	0.909	0.000571	230.0
...defend your data analysis to a team of expert physicists? — 100-level	0.443	0.715	2.99e-06	8097.0
...defend your data analysis to a team of expert physicists? — 200-level	0.875	0.842	0.00209	273.5

Figure 4.5: PhIS self-efficacy findings for all students who completed both the pre- and post-surveys. 100-level data ($N_{100} = 225$) are in blue, 200-level data ($N_{200} = 35$) in red. Each arrow represents the difference between the interpolated median of the pre-survey (tail) and post-survey (head) responses for the associated population-item.

of students reported a higher score in the post-survey than in the pre-survey. However, this metric does not account for ties, where a student gives the same response for an item in the pre- and post-surveys; ties are ignored in the calculation of the effect size. The common-language effect size can range from 0 to 1 for a given item, where 0 indicates that all respondents gave an equal or lower score to the item on the post-survey than on the pre-survey, 0.5 indicates that as many respondents reported a higher score as reported a lower one, and 1 indicates that every respondent's post-survey score was equal to or higher than their pre-survey score.

Self-efficacy

As shown in Figure 4.5, positive shifts are observed for both populations in all self-efficacy items. The figure shows interpolated medians for each item before and after the intervention; each arrow's tail represents the pre-intervention median, and each head the post-intervention median. Thus, the



	Shift	Effect size f	p-value	T_+
...design an experiment to answer a scientific question in physics?	0.899	0.835	6.78e-08	2842.0
	0.794	0.874	1.45e-06	1250.0
...look at data that you collect and characterize its patterns mathematically?	0.603	0.733	4.31e-05	3414.5
	0.490	0.686	0.0157	981.0
...understand the kinds of problems that experimental physicists would investigate?	0.659	0.737	5.06e-05	3151.0
	0.547	0.826	2.44e-05	1138.0
...contribute to a team of physicists investigating an experimental physics problem?	0.771	0.733	5.28e-05	3411.0
	0.244	0.638	0.0491	1286.0
...defend your data analysis to a team of expert physicists?	0.554	0.763	9.42e-06	3124.5
	0.232	0.624	0.0915	1068.5

Figure 4.6: PhIS self-efficacy findings for 100-level students who completed both the pre- and post-surveys, divided by gender. Data from male-identifying students ($N_M = 135$) are in dark blue, female-identifying students ($N_F = 88$) in light blue. Each arrow represents the difference between the interpolated median of the pre-survey (tail) and post-survey (head) responses for the associated population-item.

length represents pre-post change. The 100-level and 200-level populations' data are shown in blue and red, respectively. The effect sizes vary, but there is a positive shift for every self-efficacy item at a 99.7% confidence level or better ($p < 0.003$) in both populations.

Comparing the size of the populations' arrows for each item, we note that the shift in the 200-level students' responses is consistently greater than that of the 100-level students, by roughly 50% on average. This magnified effect is observed despite the 200-level pre-data medians being consistently higher than those of the 100-level students, leaving less room for improvement.

Every individual 200-level response to the item "How confident are you that you can design an experiment to answer a scientific question in physics?" either did not change or became more expertlike. This is a useful example for understanding the common-language effect size. Of all the students whose scores changed, 100% of them increased. Therefore, $f = 1$.

The PhIS results are unique among our dataset in that there are notable differences between responses given by male- and female-identifying students, as shown in Figure 4.6. While there are positive shifts for all items for both genders, female-identifying students consistently started lower on each item and saw a smaller increase. Female-identifying students' increases on the last two items are notably small; the shift in responses to **PhIS4** barely meets the traditional significance threshold $p < 0.05$, and the shift for **PhIS5** does not.

4.5.3 RQ3: To what extent are students productively engaged in the NOMR activities, and how does that engagement compare with the hands-on labs in the same course?

Following the analysis methods of Karelina *et al.* [29] described in Section 4.3.2, we produced flow plots for each week of each population's lab activities. The number of students who responded with each (x, y) pair, representing (skill/knowledge, challenge) is indicated by the size of the dot at that point. The *area* of each dot is proportional to the number of students it represents. For example, suppose 4 students responded that the lab was extremely challenging ($y = 7$) and they felt only moderately skillful ($x = 4$), and 16 students responded that the lab was a significant challenge ($y = 5$) that they felt prepared to tackle ($x = 5$). We would see a dot at $(4, 7)$ with radius R and a second dot at $(5, 5)$ with radius $2R$. The absence of a dot indicates that zero students gave the

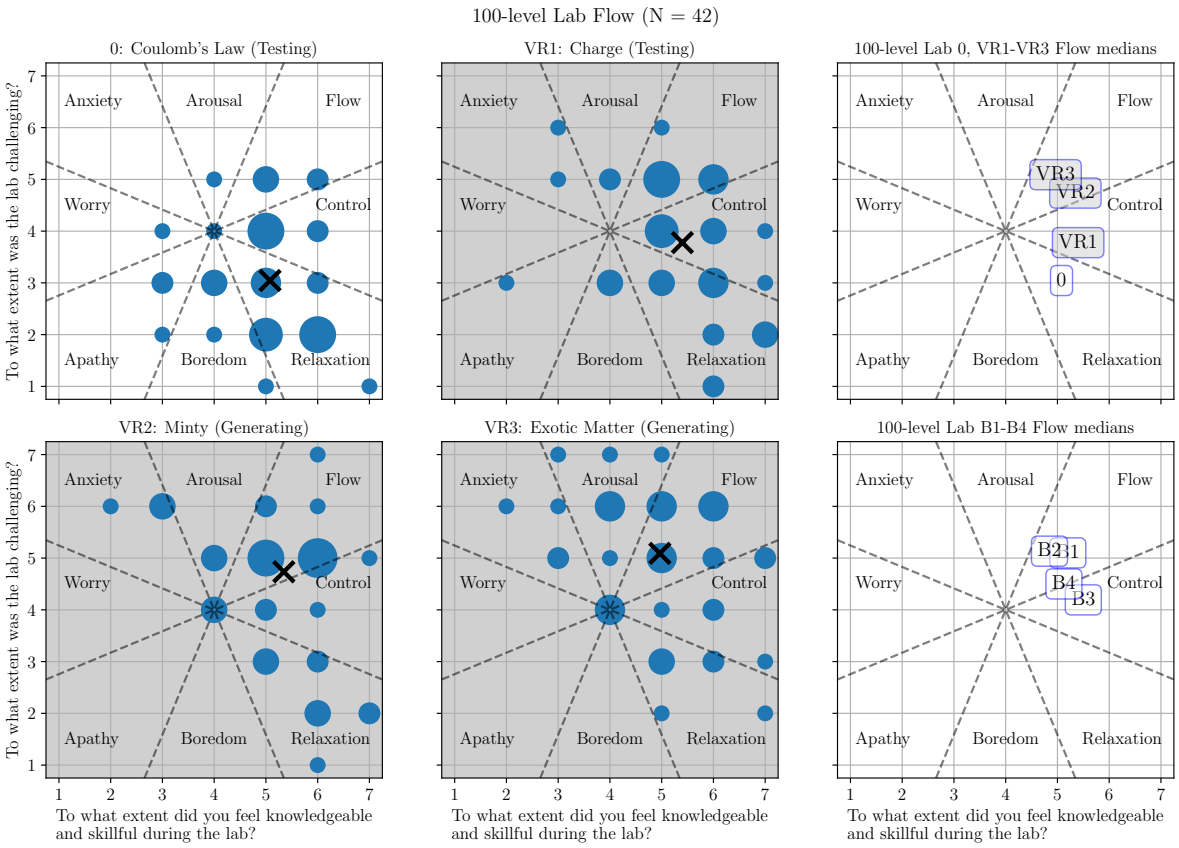


Figure 4.7: Flow plots for all $N_{100} = 42$ 100-level students who completed all eight flow surveys. Shaded plots represent VR labs. The black cross on each plot represents the 2D interpolated median of those responses. The area of each dot is proportional to the number of students it represents. The upper right and lower right plots show the medians from the first four and last four labs of the quarter, respectively.

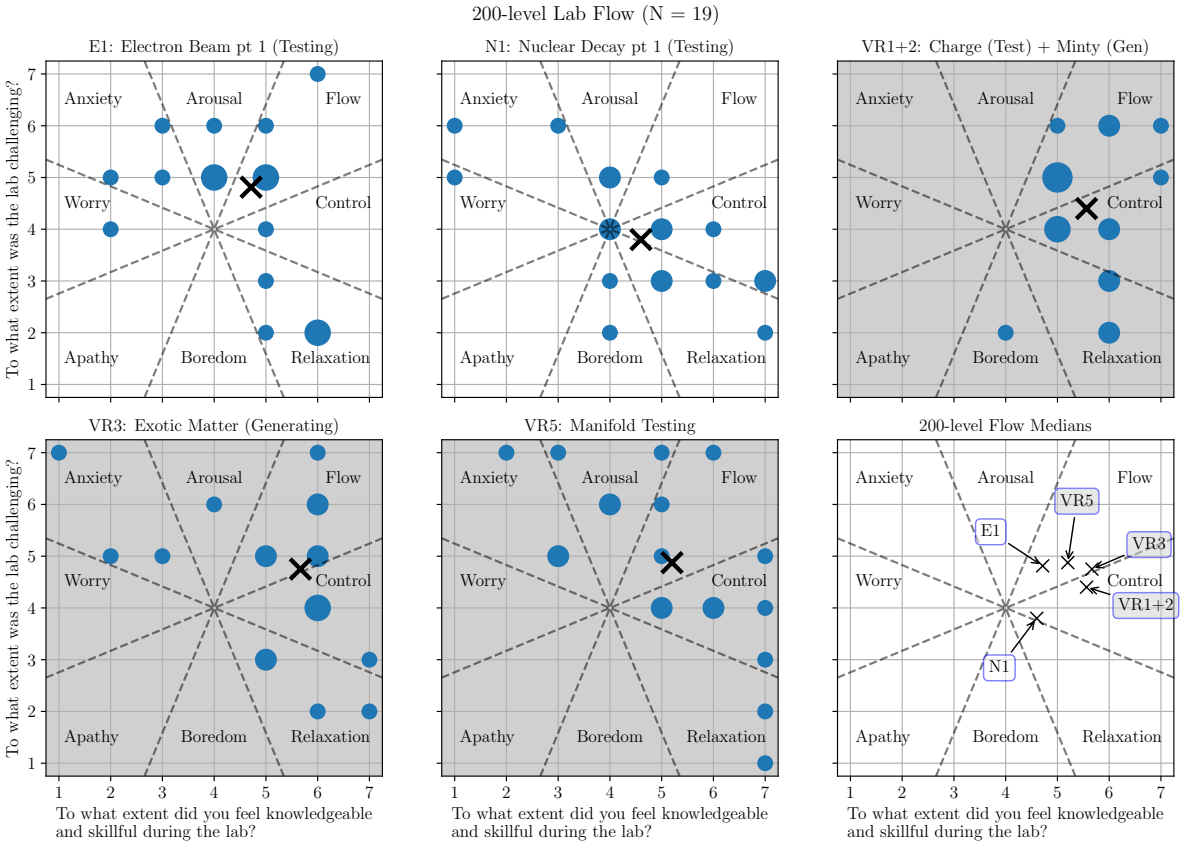


Figure 4.8: Flow plots for all $N_{200} = 19$ 200-level students who completed every flow survey. The sixth (lower right) plot shows the medians from the other five plots all together. Shaded plots and labels represent VR labs. The black cross on each flow plot represents the interpolated median of those responses. The area of each dot is proportional to the number of students it represents. Lab activities focusing on communication and writing with no new experimentation, labeled "C" in Table 4.3, are omitted.

associated response.

For our analysis, we invent a quantity to help characterize the quality of student engagement through the lens of flow across an entire class. We calculate the interpolated median of these data along each dimension of the flow plot (skill/knowledge on the x-axis, challenge on the y-axis) and plot the point defined by these medians. As a metric for the aggregate engagement state of the class, the closer to the top-right corner this 2-D median is, the more effective the lab was at inducing productive engagement. We use the interpolated median rather than the mean, as it better captures the distribution of these ordinal data.

One can produce error bars for the interpolated medians by taking the standard error along each dimension. We found this to consistently produce error bars the same size or smaller than the marker, so we have omitted them.

The 100-level population's flow data from labs 0-VR3 (the first half of the course) are shown in Figure 4.7 alongside summaries of the 2D medians for labs 0-VR3 and B1-B4. The 200-level population's flow data are plotted and their medians summarized in Figure 4.8.

We note a few key findings from the flow plots of 100-level students (Figure 4.7):

1. In both 100-level model-generating VR labs (VR2 and VR3), zero respondents reported Boredom, Worry, or Apathy. This is not the case in any other 100-level labs.
2. There is gradual migration out of Relaxation from labs 0-VR3.
3. The especially high-challenge, high-skill point in the Flow channel $(6, 6)$ had zero respondents in the first two labs 0 and VR1, one respondent in lab VR2, and several respondents in lab VR3.

Looking at the migration in the 2D medians of the 100-level students' responses to labs 0-VR3 (Figure 4.7, top right), we see a trend consistent with the design of the curriculum: it starts out easy and becomes more difficult by the week (upward movement), but students report feeling equipped to handle the increased difficulty (remain on the right side). Lab VR3, the third and final week of NOMR lab activities, had the median closest to the top-right of the plot out of all of the VR labs.

The rightmost column of Figure 4.7 lets us compare the 2-D medians of the NOMR labs to the ISLE-based hands-on circuit labs B1-B4, shown on the top-right and bottom-right of the figure,

respectively. The students complete all A labs before embarking on the B labs, which comprise the latter half of the course. We note that lab VR3 and the first two weeks of circuit labs (B1 and B2) all achieved similar states of productive engagement, landing near the diagonal in the Flow channel.

In the 200-level population, $N_{200} = 19$ students completed every survey over the course of the quarter; the flow plots and aggregated medians are shown in Figure 4.8. Lab activities focusing on communication and writing with no new experimentation (Table 4.3), are omitted from the flow plots. As with the 100-level population, the VR labs' medians follow a path up and to the left over time, starting in Control and landing solidly in Flow. The median corresponding to the final VR lab (VR5) is farthest along the diagonal bisecting the Flow channel out of all the labs. We note that only 3 student responses for VR5 are actually in the Flow region, with the majority sitting in Control and Arousal and a few in Anxiety and Relaxation. The Charge and Mint lab (VR1+2) and Exotic Matter lab (VR3) had the most individual responses in the Flow region at 7 out of 19.

Figure 4.9 shows both populations' responses to **F7** ("To what extent was the lab fun and interesting?") for each lab activity. At the 100-level, students' responses are very high for VR1, remain elevated for VR2, and their responses for VR3 are indistinguishable from any other activity.

We see a similar pattern in the 200-level students' responses to the three VR labs in their curriculum: The Charge and Mint lab (VR1+2) was seen as the most fun (median $\sim 6.7/7$), followed by the subsequent Manifold lab part A (VR3) with a half-point lower median, and the final Manifold lab part C (VR5) landed between the hands-on labs with a median of $\sim 5.5/7$.

Both populations reported the highest flow state in the last week of NOMR activities. The last NOMR lab for 100-level students was a model generating experiment (VR3). For 200-level students, it was a test of models generated in a prior experiment (VR5).

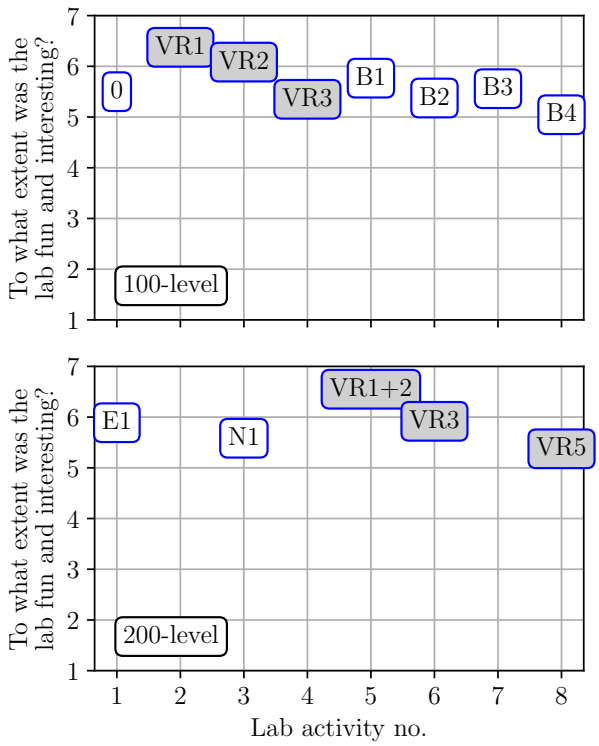


Figure 4.9: The top and bottom plots show the 100-level ($N_{100} = 42$) and 200-level ($N_{200} = 19$) populations' interpolated median response each week to **F7** ("To what extent was the lab fun and interesting?"), respectively. NOMR labs are shaded. Lab activities focusing on peer review and writing in the 200-level class are omitted.

4.6 Discussion

4.6.1 RQ1: What changes are observed in students' epistemology about experimental physics as a result of the NOMR labs?

LES1: Why are experiments a common part of physics classes?

The data show (Figure 4.3) that for both populations in our study, the frequency of each code underwent a shift consistent with growth toward an expertlike understanding of the role of experimentation in physics classes. Both populations started with the *Modeling* code occurring much less frequently than *Model testing*. After the intervention, both received the *Modeling* code more frequently and the 100-level population received *Model testing* less frequently than when they started. In particular, the 200-level population closed the gap entirely: Their post-survey frequencies for *Modeling* and *Model testing* had equal values at the end of the quarter. Movement toward equal emphasis on these two codes is considered expertlike as it is in alignment with the epistemological basis of the ISLE approach, itself grounded in the real-world practice of physics.

Supplemental learning came up much more frequently than *Scientific abilities* in pre- and post-surveys in both populations. However, both populations narrowed that gap: Both populations reported *Scientific abilities* more often and the 200-level population reported *Supplemental learning* less often than when they began, which are shifts to more expertlike belief.

The small 100-level *Supplemental learning* shift is a little surprising in context of the model-generating NOMR labs, in which students are coming up with models for completely fictitious phenomena in a clear and deliberate departure from lecture content. *Supplemental learning* is not a goal of those labs. It could be that the high and unaffected *Supplemental learning* code frequency is due to students' long history of traditional science labs where supplemental learning is indeed the primary goal; years of conditioning are not readily overcome by a 3-4 week intervention.

In sum, The LES code frequencies in both populations reveal that the students' beliefs are changed by the NOMR labs in significant ways. Regarding experimentation as part of their course-taking, they are less likely to consider classroom laboratory activities to be a supplement to the

theory they learn in lecture (primarily in the form of testing theories they have already learned), and more likely to see them as an opportunity to gain new knowledge in the form of developing scientific abilities and developing scientific models.

LES2: Why do scientists do experiments in professional research?

We compare our *Discoverment* code frequencies to the original study's findings by adding together each of their populations' *Model Development* and *Discovery* code frequencies. We expect that a subset of responses in each population received both codes, so these combined *Discoverment* code frequencies are likely over-estimates.

We would expect a collection of expert responses to **LES2** to receive both *Discoverment* and *Model testing* codes at a fairly high frequency. The original study's PhD student responses merited *Discoverment* at 65% frequency and *Model testing* at 90% [25]. These were the greatest and smallest frequencies among all populations in the original study, respectively. PhD students were the most experienced population in the original study, suggesting that expertlike change would manifest as an increase in the *Discoverment* code frequency and a decrease in the *Model testing* code frequency.

The increase in *Discoverment* frequencies in both UW populations suggest growth toward an expertlike understanding of the multi-faceted role of experimentation in scientific research. On account of the low starting point for both populations, we suggest that we can interpret any increase as developing more toward expertlike beliefs.

The *Model testing* shifts are more ambiguous: the increase in 100-level responses is opposite to what we would consider an expertlike change. However, that increase brings the 100-level *Model testing* frequency almost exactly in line with that of the PhD student population in the original study. On its own, this increase could be generously interpreted to suggest that 100-level students' belief that experiments in scientific research serve to test, support, or prove models was unchanged given the inherent uncertainty in qualitative analysis. More conservatively, this incomplete but accurate belief was bolstered alongside beliefs (i.e. *Discoverment*) that lead to a more complete expertlike understanding. The 200-level population saw a decrease in *Model testing* codes, moving

from 100% to 82% frequency. This is below that of any population in the original study. Unlike the 100-level activities, the 200-level activities–especially the Manifold Lab–are routinely connected to examples of real-world research as part of the introduction for each lab. Explicitly drawing these parallels may have contributed to the relatively large shift in the 200-level responses.

Taken as a whole, these findings represent growth toward an expertlike understanding of the role of experimentation in physics. After NOMR, students shift away from viewing experimental physics exclusively as a theory-testing endeavor, to one that includes a variety of important aspects of the role of experimentation in generating new knowledge. This shift brings students closer to to the expert view of scientific knowledge as a process that involves rigorous validation in the natural world.

4.6.2 RQ2: What changes are observed in students' physics self-efficacy in experimental physics as a result of the NOMR labs?

The presence of a positive change at a 99.7% confidence level or better ($p < 0.003$) for every PhIS self-efficacy item for both populations suggests that students' self-efficacy around conducting physics experiments is tangibly improved after participating in NOMR labs. Students believe that they are learning in ways consistent with widely agreed-upon undergraduate physics laboratory learning goals [31]. That these shifts are observed in the 200-level population is not especially surprising, as each item represents a core learning outcome for a quarter-long course for physics majors that specifically focuses on experimental physics. It is surprising that we see similar shifts in the 100-level population after just a few weeks of NOMR laboratory exercises. Still, the 100-level shifts' lesser magnitude is consistent with the relatively light depth and duration of the 100-level intervention compared with the 200-level version. All told, students' responses moved closer to those of expert physicists and indicate that their confidence in their own ability to do experimental physics is strengthened significantly.

Looking at the differences by gender in the 100-level data, the lower starting point for female-identifying students across all items is consistent with research into identity and belongingness in introductory physics [22]. Put colloquially, the field of physics is commonly thought of as being an

old boys' club, and female-identifying students have on average a harder time developing science identity in physics.

The dramatically smaller shifts in female-identifying students' responses to **PhIS4** and **PhIS5** are of particular interest. These items are focused on one's extrinsic interactions with a group of physicists than on one's intrinsic ability to perform a category of tasks. For that reason, it is plausible to believe that these items are inherently gendered; that is, administering these two items would elicit a similar difference by gender in any context. It may be the case that male-identifying students build more confidence in NOMR labs than female-identifying students do, but the absence of gender-distinct results in the LES and flow data suggest the interaction with the headsets seems to be gender neutral. Thus, we hesitate to attribute the results from these items to the instrument or the intervention, and highlight this as an area for future study.

Both populations' self-efficacy about designing, conducting, and interpreting experiments is significantly improved after working through NOMR labs. These shifts are aligned with the AAPT laboratory learning objectives [31]. We suggest these data indicate that the NOMR labs are helping students develop confidence in their professional capacity as experimentalists, while also helping them develop more expertlike habits of mind about experimental physics.

4.6.3 RQ3: To what extent are students productively engaged in the NOMR activities, and how does that engagement compare with the hands-on labs in the same course?

We see the majority of students in the productive zones of Flow, Arousal, and Control during the model-generating NOMR labs. We interpret being in a flow state as optimized student engagement in the learning activities. Achieving flow requires that students know what to do, how to do it, and how well they are doing; students tuning out or becoming lost in the face of the open-endedness of the activities would be reflected in low-skill responses on the left of the flow plots. The data show that the NOMR labs are providing just enough scaffolding to keep students in the zone of proximal development and in flow [41]. Of the first series of 100-level labs (0-VR3), VR3 induced the most productive aggregate state of engagement in students; no responses indicated a state of Worry, Apathy, or Boredom. The 200-level population achieved more productive engagement with

the VR labs than either of the hands-on labs E1 and N1.

We recognize that VR is an engaging environment on its own. While it may be hard to disentangle the novelty of VR from the activities themselves, we do see evidence that the novelty wears off. We consider the effect of the gaming/entertainment appeal of VR by examining student responses to item **F7** (plotted in Figure 4.9): "To what extent was the lab fun and interesting?" The novelty effect is associated with the introduction of an exciting new technology in the classroom, which induces an initial boost in student engagement that eventually wanes [9].

We estimate that VR's novelty lasts two weeks in our context, as we observe in both populations the highest **F7** score in the first NOMR lab, the second highest in the second NOMR lab, and the third NOMR lab is no more or less fun than any of the hands-on labs in the course.

Despite the third week of NOMR labs not benefiting from the novelty effect, students reported the greatest aggregate flow state during that activity (VR3 and VR5 at the 100-level and 200-level, respectively). This suggests that the novelty effect does not fully explain the productive and deep engagement with VR physics labs. Students are not engaged simply because VR is fun; they are engaged because the physics is compelling. NOMR labs use VR specifically because its "secret sauce" of hands-on interaction with fictitious physical phenomena is otherwise impossible. We consider students' strong engagement after the novelty has worn off as evidence that NOMR labs may be leveraging the unique affordances of VR in a pedagogically useful way.

Our comparison of VR and hands-on labs contrasts with Karelina *et al.*'s comparison [29] between students' engagement with video labs and hands-on labs. They found students reported video labs to be slightly more challenging, less fun, and that they felt less skillful when compared to hands-on labs. Our findings demonstrate that students' engagement with VR labs can be similar to or better than their engagement with hands-on labs in the same course. It is important to note that Karelina *et al.* compared two distinct populations of students who went through hands-on and video versions of the same lab activity, while our study compares responses to different lab activities from the same population, so comparisons between our findings and theirs should be made with caution.

We hesitate to overinterpret our analysis of the flow data. In this study we analyze absolute

rather than relative scores. The original application of the eight-channel flow model [34] collected responses from each participant at many points in time over several days. The researchers determined each participant's average response for an item. When creating flow plots, they plotted z-scores relative to each participant's average response to each item. This method accounts for the fact that every individual interprets Likert scale questions differently; one person's 3/7 is another's 6/7. Karelina *et al.* [29] adapted this original methodology in favor of examining absolute scores due to the limitations of a classroom setting. They had no more than 2-3 responses from any given participant, and we replicated their methodology for comparability.

Further, we note that flow states manifest in neurodiverse learners in ways that are not fully understood [37]. The fact that the measurements of students' flow state takes place in a group learning context adds the complexities of group social interactions to the picture. Flow is an individual measurement of an experience that occurs in a group context, which does not give us any information about group dynamics and cohesion. These shortcomings are excellent areas for further work.

4.6.4 Use of fictitious physics with virtual reality

Existing survey data does not fully capture the in-class and in-headset experience of NOMR labs. Students' experimental results in NOMR labs are different from reexaminations of well-understood phenomena: Every group's findings are new knowledge students have generated from scratch. We posit that, in this way, NOMR labs may allow students to experience the satisfaction of discovery that professional physicists find so compelling, as reflected in feedback from post-course surveys:

> "I was thrilled and enlightened to be put in a position to analyze physical phenomena that were undocumented and that I had never heard of. Being able to work with a pocket universe and using experimentation to describe it was the best experience in physics courses I've ever experienced. I preferred this VR experience to any physical lab for the sole reason of it being entirely new and having to get every ounce of info about it through experimentation and collaboration."

> "It's been so much fun learning physics in an exploratory way that focuses on letting us be creative with our thinking. I've not only learned a lot about error analysis and creating models, but also gained a much better perspective on how science and research 'work' in the real world."

As NOMR lab instructors and experienced teachers, we observe students engaging in scientific creativity in ways we have not seen before. Rigorous characterization of creativity is a challenging research task, and is not captured in the surveys we administered. We postulate that the use of VR is in part responsible for helping unleash student creativity, and highlight this as an area for further study.

This study does not create evidence one way or another about whether VR is necessary for the implementation of the fictitious physics approach to stimulate creativity in introductory labs. We predict that such an intervention would not be as effective without VR, in that the joy of discovery expressed above would likely be altered by the difficulties associated with non-immersive simulations. Manipulating objects in 3D space on a 2D screen can be challenging, not to mention that it risks widening the conceptual gap between the novel physics and the physics of our universe. Being overly disconnected from a physically interactive 3D space may preclude the suspension of disbelief that allows students to engage so readily in NOMR. It is our experience that the preceding argument is difficult to make convincingly without sharing the in-headset experience; making it rigorously will require substantial human-computer interaction research.

There remains a possibility that a non-VR version's effectiveness could be sufficient for a headset-free version of NOMR to be a favorable trade-off, considering the expense and overhead associated with VR technology. Presently, Meta Quest 2 headsets (the same used for this study) cost $\$300$ each; at eight lab groups to a section, the cost of outfitting a classroom to run NOMR labs is $\sim \$2400$. While not outrageous, this is beyond the reach of many physics laboratory budgets, especially in under-served communities. Phone-based virtual or augmented reality, 3D simulations experienced on a monitor and controlled via mouse and keyboard, or entirely 2D simulations could all employ fictitious physics in a similar way to NOMR labs without the expense of full VR

headsets. Further developments based on the work in this paper can contribute to exploring these avenues to open up broader access to this instructional innovation.

Research Question	Survey	Outcome
RQ1: What changes are observed in students' epistemology about experimental physics as a result of the NOMR labs?	LES	Students become more expertlike in their epistemologies associated with the role of experimentation in learning, and in research. NOMR students shift away from viewing experimental physics exclusively as a theory-testing endeavor, to one that includes a variety of important aspects of the role of experimentation in generating new knowledge.
RQ2: What changes are observed in students' physics self-efficacy in experimental physics as a result of the NOMR labs?	PhIS	NOMR labs help students develop belief in their professional capacity as experimentalists, while also helping them develop more expertlike habits of mind about experimental physics.
RQ3: To what extent are students productively engaged in the NOMR activities, and how does that engagement compare with the hands-on labs in the same course?	FS	Students become increasingly engaged with successive NOMR labs, even after the novelty wears off. They are most engaged when developing their own hypotheses with novel physics.

Table 4.4: The research questions, surveys, and outcomes of this study are summarized.

4.6.5 Epistemological hazards of fictitious physics

The use of fictitious laws of physics raises concerns about whether interacting with fictitious laws of physics can negatively affect students' physical intuition and conceptual understanding. These concerns have been on our minds since the first trial of NOMR in early 2020. For that reason, we take care to maintain conceptual separation between NOMR-unique physics and the physics of our universe:

- The introduction of every NOMR lab manual (except VR1: Charge) makes clear that the physics students will be investigating was created for the purpose of the lab and does not exist in our universe. The lab manuals frame activities with fictitious physics as being original investigations into unknown physics, putting students in the shoes of Coulomb and peers.
- Fictitious particles are given silly names (e.g. minty particles); where they are not named in the lab manual, students are encouraged to come up with their own names for the particles. They discovered them, after all.
- At no point do students work with simulations of both real and fictitious physics at the same time.
- Fictitious particles are only ever referenced in context of the laboratory component of a course; they are not mentioned in lecture or tutorial components.

To date, we have not seen evidence of negative impacts on students' conceptual or procedural physics knowledge arising from their work with fictitious physics.

4.7 Conclusions

Steering students away from confirmation of known facts and into a different (simulated) universe sounds like a day in the life of Ms. Frizzle's class [10]; in lieu of a magic school bus, we use VR headsets. In both cases, students are transported to the teacher's choice of hands-on learning environment to create knowledge through collaboration with their peers in a fun, engaging, and memorable environment. In NOMR, students learn to gain knowledge in the way that experts do.

This study contributes to the physics education community's effort to create laboratory activi-

ties that foster students' growth along affective and epistemological dimensions. We have demonstrated that these goals can be achieved by inventing new physics for students to explore, effectively drawing a new frontier of physics at the introductory level. Our findings (summarized in Table 4.4) begin to paint a positive picture of the affective impacts of labs featuring fictitious physics, but our experience as instructors suggests that we have yet to capture the full effect of this approach.

Subsequent analysis will include respondent-relative analysis of flow data and examine students' responses across surveys to identify correlations between change toward expertlike beliefs, engagement in lab activities, and development of self-efficacy.

That student engagement remained strong despite the decay of the novelty of VR is an important finding to consider for others interested in the role of immersive technologies in physics education. Replicating traditional or 2D simulator-based physics instructional activities in VR has a track record of being more engaging than comparable treatments in other media, but yielding little to no educational benefit in comparison [51, 40, 33]. Our findings suggest that instead, educators, researchers, and developers interested in the use of immersive technology in the physics classroom should collaborate to identify niches where its affordances can be leveraged to create unique learning experiences.

Whether VR is strictly necessary to implement the fictitious-physics approach is difficult to say with scientific rigor, and while we do not attempt to do so here, we predict that it is necessary to a greater or lesser degree. That this study found positive results using a VR intervention suggests that the utility of VR in physics education lies in niches where it lets us create learning experiences that would be infeasible or impossible by other means. In other words, an effective educational VR activity is not effective because it uses VR any more than a PhET [39] is effective because it runs in a browser. Rather, it creates a learning experience tailored to solve a specific pedagogical problem, using the medium that provides the most appropriate foundation for doing so.

4.8 Acknowledgements

The authors thank Sabrina Cheng and Tyler Flom for their assistance coding the LES data, Kazumi Tolich for major contributions to NOMR's instructional materials and allowing us to administer

all of these surveys in her course, David Aplin and Eddie Mendoza for heroically taking on all of the technical legwork to facilitate NOMR's classroom implementation, and Eugenia Etkina, Peter Shaffer, and Charlotte Zimmerman for their insightful feedback on early drafts of this article. This work would not have been possible without the many TAs who have taught NOMR labs and the supportive community of the UW Physics Education Research Group; both groups are too numerous to list here, but you know who you are. Thank you.

We gratefully acknowledge financial support from the UW Student Technology Fee committee and the National Science Foundation GRF (DGE-1256082).

Chapter 5

CONCLUSION

Learning to create and validate new knowledge through an empirical lens is a critical step in the development of budding scientists, and a highly desirable one in members of an informed public. This dissertation represents the genesis of a novel paradigm in physics laboratory instruction that creates a new frontier of physics for students to develop their capacity as creators of knowledge.

Guiding students away from a confirmatory mindset and developing their belief in themselves as sources of authority about experiments in physics is a prominent obstacle to epistemologically authentic physics laboratory instruction. A confirmatory mindset is easily reinforced by the presence of a "correct" answer, which are in ample supply via the Internet. In Chapter 2, I discuss the idea of creating a learning environment where where are no "correct" answers to be found. That idea combines with other design principles to form the pedagogical basis of NOMR, upon which a lab curriculum leveraging several real and fictitious systems of physical phenomena is built.

Chapter 3 follows a chronological account of the development of NOMR with a focus on the development of the VLE software, classroom implementation practices, and instructional materials comprising NOMR. NOMR's precursors in the realms of special relativity and Gauss' law provide the foundation for the first prototype and pilot test of NOMR. Classroom safety procedures, VR locomotion and interaction design patterns, and technical implementation of the NOMR physics engine are covered in detail. Lessons learned from each implementation of NOMR give way to a stable routine of classroom implementation independent of my involvement, and my focus turns to education research on NOMR's effects.

Studies of NOMR's affective effects are described in Chapter 4, itself a standalone paper soon to be published in Physical Review: Physics Education Research. NOMR's effects are assessed along axes of students' epistemology about experimental physics, self-efficacy in experimental

physics contexts, and engagement quality with each lab exercise. Gains are observed along all three axes.

The aforementioned studies do not capture the robust, energetic creativity we see students engage in with NOMR labs. In my and my colleagues' experience as instructors, such creativity is nigh unheard-of in other physics lab contexts. Characterizing this creativity is one of many research questions about NOMR's effects which merit further exploration. The process of taking steps toward this characterization demonstrated NOMR's efficacy and produced yet more questions:

1. How does group makeup across different demographic lines affect students' experience of NOMR labs?
2. How do outcomes from NOMR labs compound across multiple courses with NOMR labs?
3. What longitudinal (long-term) effects are observed in students who completed NOMR labs, compared to students who did not?
4. Quite a lot of E-CLASS data were gathered from NOMR and non-NOMR lab courses. Are there any interesting patterns in there?
5. How can NOMR or other fictitious-phenomena approaches be used to develop skills associated with successful careers in physics theory?

Beyond questions about its effects, NOMR enables the investigation of other physics education research questions:

1. What happens if we ask students to come up with field and potential functions for NOMR phenomena (e.g., minty particles?) Do students see field and potential as mathematical concepts independent of the physical phenomena (gravitation, electromagnetism) where they are introduced?
2. How do different types of readouts (decimal, measuring stick ticks, circular gauge, etc.) affect students' ideas about quantifying uncertainty for the associated quantities?

Dave Aplin, Eddie Mendoza, and the NOMR Student Developer Group are continuing development of NOMR in several promising directions:

1. Real-time numerical simulation of conductors of various geometries;
2. A tool enabling NOMR users to change the magnitude of particles' charge;

3. A new scenario featuring a particle that decays into a number of visible daughter particles that can be analyzed to infer the properties of another, invisible daughter particle;
4. In-app tutorials and prompts.

At the time of writing, NOMR is used as part of UW-Seattle and WSU's physics laboratory curricula. Anyone interested in adding their institution to that list is warmly encouraged to contact the UW Introductory Physics Lab Manager (David Aplin[1] at the time of writing) and myself.[2]

My work has shown that NOMR is feasible, usable, and works as intended. I aspire to think of this milestone as the end of the beginning for NOMR, which is now ready for the next generation of developments, questions, and scholars to ask them.

Want to learn more? This dissertation describes the development and evaluation of an interactive virtual reality learning environment. Capturing the in-headset experience is impossible with static text and images. At the following permanent URL, one can find links to video demonstrations of this project, its current web presence, contact information, and links to any public version which may later be published, which will be updated as the project continues to evolve.

https://purl.org/nomr/dissertation

---

[1] daplin@uw.edu

[2] jpcanright@gmail.com

BIBLIOGRAPHY

[1] Wendy K Adams and Carl E Wieman. Development and Validation of Instruments to Measure Learning of Expert-Like Thinking. *International Journal of Science Education*, 33(9):1289–1312, 2011.

[2] Saalih Allie, Andy Buffler, Loveness Kaunda, and Margaret Inglis. Writing-intensive physics laboratory reports: Tasks and assessment. *The Physics Teacher*, 35(7):399–405, 10 1997.

[3] Albert Bandura. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2):191–215, 1977.

[4] P. A. Bartlett and K. Dunnett. Secret objectives: promoting inquiry and tackling preconceptions in teaching laboratories. 5 2019.

[5] David T Brookes, Eugenia Etkina, and Gorazd Planinsic. Implementing an epistemologically authentic approach to student-centered inquiry learning. *Physical Review Physics Education Research*, 16(2):20148, 12 2020.

[6] David T. Brookes, Eugenia Etkina, and Gorazd Planinsic. Implementing an epistemologically authentic approach to student-centered inquiry learning. *Physical Review Physics Education Research*, 16(2), 12 2020.

[7] Jared P. Canright, Jack R. Olsen, and Suzanne White Brahmia. Leveraging virtual reality for student development of force models in the introductory lab. In *Physics Education Research Conference*, pages 75–80. American Association of Physics Teachers (AAPT), 9 2020.

[8] Jared P. Canright and Suzanne White Brahmia. Developing expertlike epistemologies about physics empirical discovery using virtual reality. pages 75–80. American Association of Physics Teachers (AAPT), 10 2021.

[9] Richard E Clark. Reconsidering research on learning from media. *Review of educational research*, 53(4):445–459, 1983.

[10] Joanna Cole and Bruce Degen. *The Magic School Bus Lost in the Solar System*. Scholastic, 10 1990.

[11] Mihaly Csikszentmihalyi. Flow: The Psychology of Optimal Experience. 1 1990.

[12] Mihaly Csikszentmihalyi. *Applications of flow in human development and education: The collected works of Mihaly Csikszentmihalyi.* Springer Science + Business Media, New York, NY, US, 2014.

[13] James Day, Jared B Stang, N G Holmes, Dhaneesh Kumar, and D A Bonn. Gender gaps and gendered action in a first-year physics laboratory. *Physical Review Physics Education Research*, 12(2):20104, 2016.

[14] E. Etkina, D. T. Brookes, and G. Planinsic. The Investigative Science Learning Environment (ISLE) approach to learning physics. In *Journal of Physics: Conference Series*, volume 1882. IOP Publishing Ltd, 5 2021.

[15] Eugenia Etkina. Millikan award lecture: Students of physics - Listeners, observers, or collaborative participants in physics scientific practices? *American Journal of Physics*, 83(8):669–679, 8 2015.

[16] Eugenia Etkina and Alan Van Heuvelen. Investigative Science Learning Environment - A Science Process Approach to Learning Physics. In *Research-Based Reform of University Physics*, volume 1. 4 2007.

[17] Eugenia Etkina and Gorazd Planinsic. Defining and Developing 'Critical Thinking' Through Devising and Testing Multiple Explanations of the Same Phenomenon. *The Physics Teacher*, 53(7):432–437, 10 2015.

[18] Eugenia Etkina, Alan Van Heuvelen, Suzanne White-Brahmia, David T. Brookes, Michael Gentile, Sahana Murthy, David Rosengrant, and Aaron Warren. Scientific abilities and their assessment. *Physical Review Special Topics - Physics Education Research*, 2(2), 8 2006.

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

[20] Scott Freeman, Sarah L Eddy, Miles McDonough, Michelle K Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415, 6 2014.

[21] Genymobile. scrcpy Github page.

[22] Zahra Hazari, Philip M Sadler, and Gerhard Sonnert. The Science Identity of College Students: Exploring the Intersection of Gender, Race, and Ethnicity. *Source: Journal of College Science Teaching*, 42(5):82–91, 6.

[23] N. G. Holmes, Grace Heath, Katelynn Hubenig, Sophia Jeon, Z Yasemin Kalender, Emily Stump, and Eleanor C Sayre. Evaluating the role of student preference in physics lab group equity. *Physical Review Physics Education Research*, 18(1):10106, 1 2022.

[24] N. G. Holmes, Jack Olsen, James L. Thomas, and Carl E. Wieman. Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content. *Physical Review Physics Education Research*, 13(1), 5 2017.

[25] Dehui Hu and Benjamin M. Zwickl. Examining students' personal epistemology: the role of physics experiments and relation with theory. In *PERC Proceedings*, pages 11–14. American Association of Physics Teachers (AAPT), 3 2018.

[26] Dehui Hu and Benjamin M. Zwickl. Examining students' views about validity of experiments: From introductory to Ph.D. students. *Physical Review Physics Education Research*, 14(1), 4 2018.

[27] Dehui Hu, Benjamin M. Zwickl, Bethany R. Wilcox, and H. J. Lewandowski. Qualitative investigation of students' views about experimental physics. *Physical Review Physics Education Research*, 13(2), 11 2017.

[28] Susan A Jackson and Herbert W Marsh. Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *Journal of Sport and Exercise Psychology*, 18(1):17 – 35, 1996.

[29] Anna Karelina, Eugenia Etkina, Peter Bohacek, Matthew Vonk, Michael Kagan, Aaron R. Warren, and David T. Brookes. Comparing students' flow states during apparatus-based versus video-based lab activities. *European Journal of Physics*, 43(4), 7 2022.

[30] Gerd Kortemeyer, Philip Tan, and Steven Schirra. A Slower Speed of Light: Developing Intuition about Special Relativity with Games. In *FDG '13 Proceedings of the International Conference on the Foundations of Digital Games*, pages 400–402, 2013.

[31] Joseph Kozminski, Heather Lewandowski, Nancy Beverly, Steve Lindaas, Duane Deardorff, Ann Reagan, Richard Dietz, Randy Tagg, Melissa Eblen-Â-Zayas, Jeremiah Williams, Robert Hobbs, and Benjamin Zwickl. AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum. Technical report, AAPT, 2014.

[32] A S Mackenzie, I Newton, and H Cavendish. *The Laws of Gravitation: Memoirs by Newton, Bouguer and Cavendish, Together with Abstracts of Other Important Memoirs*. Scientific memoirs. American book Company, 1900.

[33] J H Madden, A S Won, J P Schuldt, B Kim, S Pandita, Y Sun, T J Stone, and N G Holmes. Virtual Reality as a Teaching Tool for Moon Phases and Beyond. In *PERC Proceedings*, Washington, DC, 2018.

[34] Fausto Massimini and Massimo Carli. *Optimal experience: Psychological studies of flow in consciousness.* Cambridge University Press, New York, NY, US, 1988.

[35] Fausto Massimini, Mihalyi Csikszentmihalyi, and Massimo Carli. The Monitoring of Optimal Experience A Tool for Psychiatric Rehabilitation. *The Journal of Nervous and Mental Disease*, 175(9), 1987.

[36] Lillian C. McDermott and Peter S. Shaffer. Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. *American Journal of Physics*, 60(11):994–1003, 11 1992.

[37] Andy McDonnell and Damian Milton. Going with the flow: reconsidering 'repetitive behaviour' through the concept of 'flow states'. In Glenys Jones and Elizabeth Hurley, editors, *Good Autism Practice: autism, happiness and wellbeing*, pages 38–47. BILD, Birmingham, UK, 10 2014.

[38] NASA. Orbits and Kepler's Laws.

[39] Katherine Perkins, Wendy Adams, Michael Dubson, Noah Finkelstein, Sam Reid, Carl Wieman, and Ron LeMaster. PhET: Interactive Simulations for Teaching and Learning Physics. *The Physics Teacher*, 44(1):18–23, 1 2006.

[40] Chris D Porter, Jonathon R Brown, Joseph R Smith, Erik M Stagar, Amber Simmons, Megan Nieberding, Abigail Ayers, and Christopher Orban. A controlled study of virtual reality in first-year magnetostatics. In *Physics Education Research Conference 2019*, PER Conference, Provo, UT, 7 2019.

[41] N Sanjay Rebello and Dean Zollman. Problem Solving and Motivation - Getting our Students in Flow. In *Physics Education Research Conference 2013*, PER Conference Invited Paper, pages 39–41, Portland, OR, 7 2013.

[42] Joseph M. Reilly, Eileen McGivney, Chris Dede, and Tina Grotzer. Assessing Science Identity Exploration in Immersive Virtual Environments: A Mixed Methods Approach. *Journal of Experimental Education*, 89(3):468–489, 2021.

[43] Holli Riebeek. Planetary Motion: The History of an Idea That Launched the Scientific Revolution, 7 2009.

[44] Anthony Russo, Joe Russo, Christopher Markus, Stephen McFeely, Stan Lee, Jack Kirby, and Kevin Feige. Avengers: Infinity War, 2018.

[45] Schell Games. I Expect You to Die for Oculus Quest, 2019.

[46] Daniel L Schwartz, John D Bransford, David Sears, and others. Efficiency and innovation in transfer. *Transfer of learning from a modern multidisciplinary perspective*, 3:1–51, 2005.

[47] Peter S. Shaffer and Lillian C. McDermott. Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of instructional strategies. *American Journal of Physics*, 60(11):1003–1013, 11 1992.

[48] Emily M. Smith and N. G. Holmes. Best practice for instructional labs. *Nature Physics*, pages 1–2, 5 2021.

[49] Emily M. Smith, Martin M. Stein, and N. G. Holmes. How expectations of confirmation influence students' experimentation decisions in introductory labs. *Physical Review Physics Education Research*, 16(1), 6 2020.

[50] Emily M Smith, Martin M Stein, and Natasha Holmes. Surprise! Shifting students away from model-verifying frames in physics labs. In *Physics Education Research Conference 2018*, PER Conference, Washington, DC, 8 2018.

[51] Joseph R. Smith, Amber Byrum, Timothy M. McCormick, Nick Young, Chris Orban, and Chris D. Porter. A controlled study of stereoscopic virtual reality in freshman electrostatics. In *PERC Proceedings*, pages 376–379, Cincinnati, OH, 2017.

[52] Martin M Stein, Emily M Smith, and N G Holmes. Confirming what we know: Understanding questionable research practices in intro physics labs. In *PERC Proceedings*, Washington, DC, 2018.

[53] Ricardo Trumper. The Physics Laboratory - A Historical Overview and Future Perspectives. *Science and Education*, 12(7):645–670, 2003.

[54] Michael Vignal, Gayle Geschwind, Benjamin Pollard, Rachel Henderson, Marcos D Caballero, and H J Lewandowski. Survey of physics reasoning on uncertainty concepts in experiments: an assessment of measurement uncertainty for introductory physics labs, 2023.

[55] Cole Walsh, Katherine N. Quinn, C. Wieman, and N. G. Holmes. Quantifying critical thinking: Development and validation of the physics lab inventory of critical thinking. *Physical Review Physics Education Research*, 15(1), 5 2019.

[56] Bethany R. Wilcox and H. J. Lewandowski. Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics. *Physical Review Physics Education Research*, 13(1), 2 2017.

[57] Frank Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80–83, 1945.

[58] V K Zaretskii. The Zone of Proximal Development. *Journal of Russian & East European Psychology*, 47(6):70–93, 11 2009.

Appendix A

EXAMPLE LAB REPORTS

Pilot-Level Grant Report

[redacted]

Introduction

We were assigned pocket A(1986) which has 1-3 particles and were given the task of exploring the relationship between these different particles and observe how they each behave around each other. We came up with testable findings and equations surrounding these particles to help distinguish them from each other.

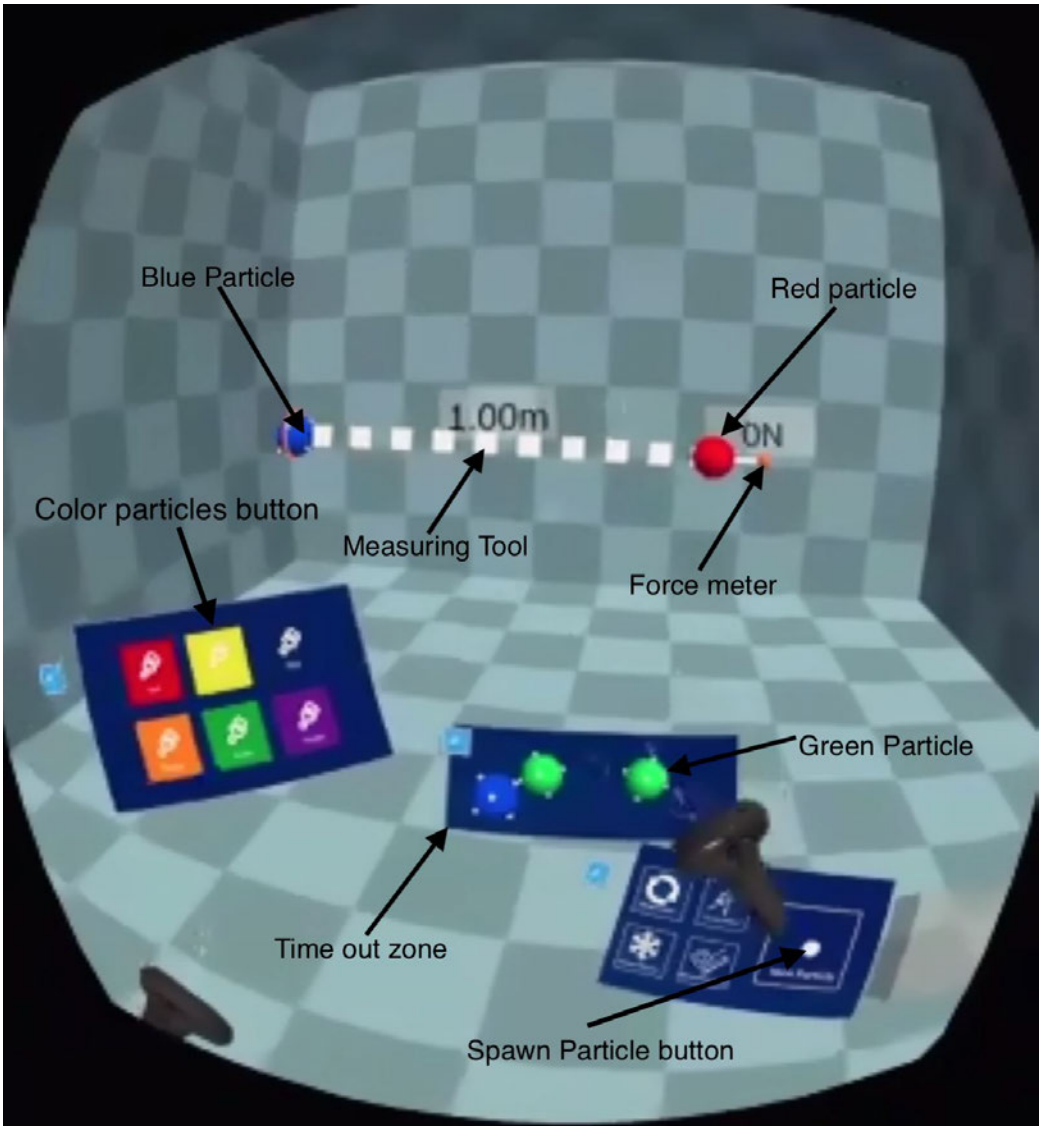
Methods

When we first entered the pocket, we started by spawning many particles to see if we noticed anything substantial in the first minute inside the pocket. We quickly noticed that there were 2 blobs of particles that attracted particles of the same type but repelled particles of the opposite blob. We also noticed a third type of particle that was repelled from itself and didn't have any other particles attached to it. We played around with this third particle and noticed it wasn't either attracted or repelled from one blob. We colored the particles in this blob red. We brought it near the other blob and observed it would attach to parts of the blob but sometimes was shot off if we tried it near other parts. We realized that the reason it was getting repelled was that the repulsion force between particles of it's type was most likely stronger than the attraction force between the other particles in the blob that repelled from red particles, something to test later on. We deleted all the particles and restarted, this time spawning in one particle at a time and labeling them with colors based on if they were attracted to each other or repelled. We went slower this time, checking the relationship between all particles before labeling them. We labeled the particle that was repelled from itself as blue. We labeled the blob that was neutral with the blue particles as red, and we labeled the other blob that attracted green particles as Blue. We wanted to get a stronger understanding of exactly what was happening between the particles.

We decided the best way to find a lot of information about the particles was to test the force between each pair of particles at different distances. In our experiment we identified distance as our independent variable and force as the dependent variable. We used the measuring tool to easily measure the distance between each particle, and used the force meter to attach one of the particles to, so we could measure the force exerted by the other particle on the one attached to the force meter.

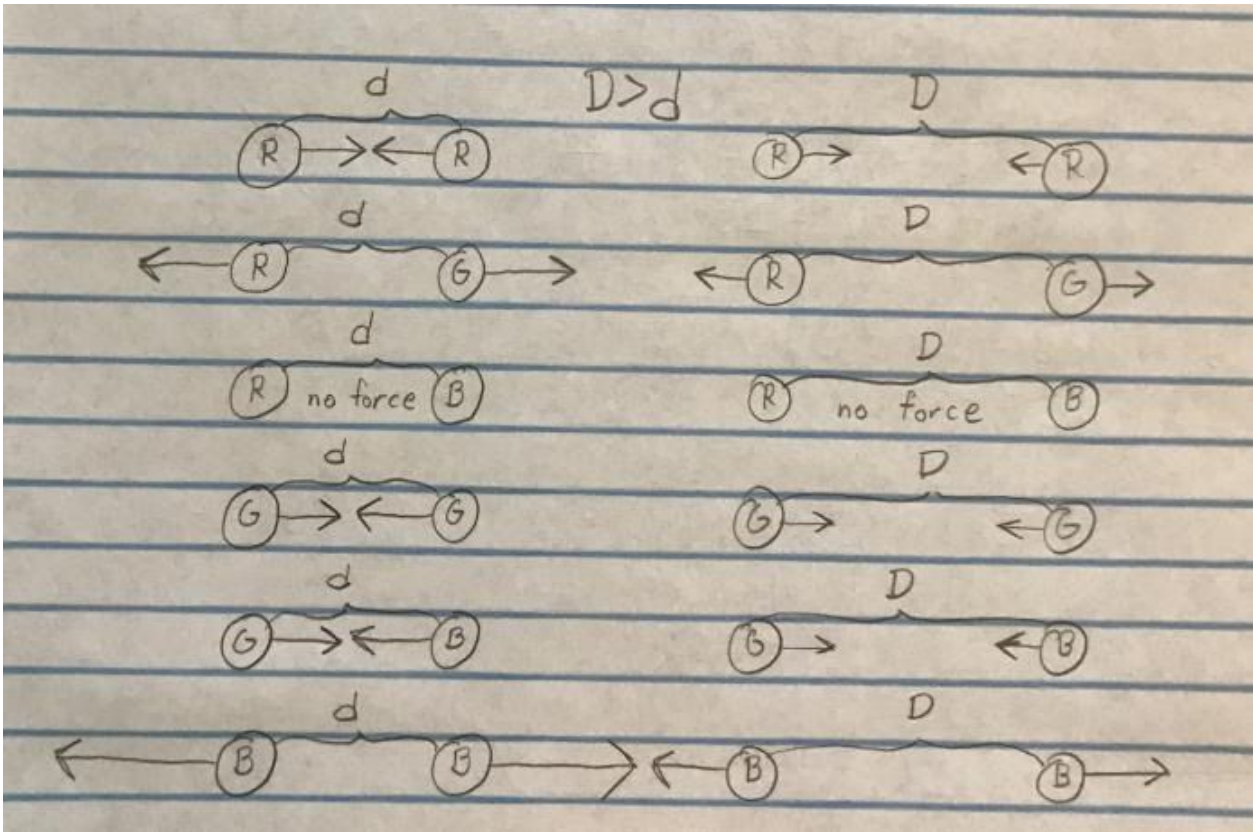
We started by moving 2 blue particles and 2 green particles into the time out zone so we could use them later, and deleted all other particles of those types. We also deleted all but 2 red particles. With the 2 remaining particles, we anchored one in place and placed the other inside a force meter. We picked out distances of 0.25m, 0.5m and 1m to measure at. We took

measurements for the relationships between all pairs of particles, as seen in the data table in the appendix. We used the following lab setup shown below.



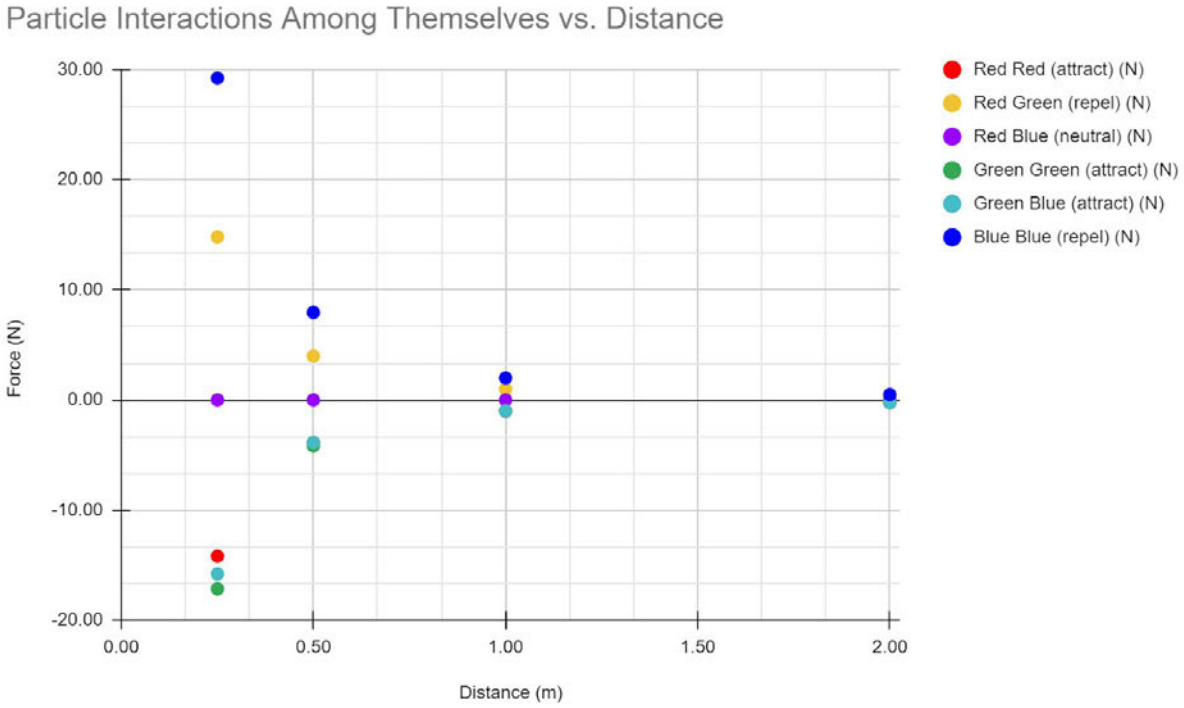
Data Analysis

In the experiment, we find that there are three different kinds of particles, and that the interactions among them vary. We labeled these particles Red, Green, and Blue. Red particles attract each other, repel Green particles, and are neutral towards Blue particles. Green particles also attract each other and, as mentioned above, repel Red particles. Blue particles are attracted to Green particles but repel each other. Each of these interactions (with the exception of Red-Blue interactions, which are neutral) are proportional to the inverse square of the distance between the particles. That is, as the distance between these particles decreases, the force they apply on each other increases. Below is a sketch to illustrate these findings.



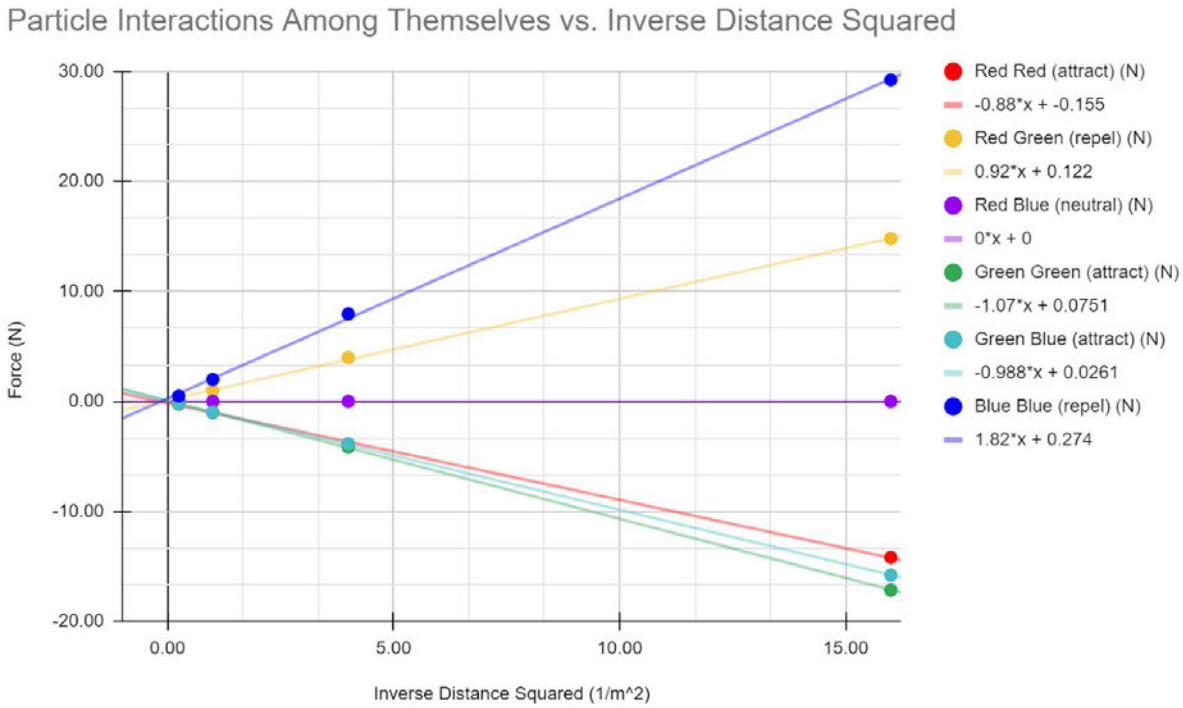
You'll notice that among these interactions, each of them have the same magnitude except for that between two Blue particles, which is twice that of the other forces. Also note that, as mentioned, Red-Blue interactions have a force magnitude of zero regardless of the distance between them.

Below is a graph of the resulting interactions, where a negative force signifies an attraction and a positive force signifies a repulsion. Once again, note that the attractive and repulsive forces have the same magnitude with the exception of Blue-Blue interactions. Another thing worth keeping in mind is that Green-Blue interactions don't have a magnitude between that of Green-Green interactions and Blue-Blue interactions—as one might expect given that Blue-Blue force magnitudes are twice that of Green-Green's—rather Green-Blue interactions are equal in magnitude to those of Green-Green.



It may be difficult to see, but the Red-Red, Green-Green, and Green-Blue interactions are indeed of equal magnitude. They are shown overlapping in this plot.

We now linearize this graph to show that the force is inversely proportional to the square of the distance between the particles.



The graph supports our statement on the proportionality and again we find the overlapping magnitudes. You may notice that as the distances get really small (so the inverse square of the distance gets really large) the Red-Red, Green-Green, and Green-Blue data points stop overlapping, but this is due to small errors in measurements (within a 10% error margin from the expected $1/x^2$ function) as the dependent force variable becomes more sensitive to the independent distance variable, not to differences in proportionalities.

Conclusion

We noticed that the values of the force looked very similar for nearly every possible interaction between the three different particles and sure enough, we found that the force between all six different combinations of particles, besides Red-Blue, where there is no force, appear to be dependent on the inverse square of the distance between the particles. Some of these forces are attractive, and some are repulsive, but the magnitude of the force for all six interactions, including the Red-Blue interaction where there doesn't appear to be a force present, can be described by one equation. Force (N) = k/(x^2), where x is the distance (m), and k is the constant of proportionality with units N(m^2). From the data we collected, we found that k can either be 0, 1 or 2 N(m^2). The value of k is equal to 0 for the Red-Blue interaction, 2 for the Blue-Blue interaction, and 1 for the other four interactions. The force being dependent on the inverse square of the distance looks quite familiar, as this same relationship is also present in

Coulomb's law, which describes the force between electrically charged particles. Although we are dealing with hypothetical particles, the similarity is worth mentioning.

One shortcoming in our experiment was that, while we did collect data from three different distances, three of the four distances were between 0.25m and 1m, with the last one being at 2m. The amount of data we collected seemed sufficient enough for us to derive equations for the particles' behavior, but we could've recorded data for more large distances, for the sake of having a more compelling and complete data set. Recording a larger range of distances could have potentially revealed patterns or results that we wouldn't have expected based on the data from the distances we did record.

This experiment was done to find mathematical representations for the behavior of three hypothetical particles, and more specifically, how much force they exert on each other, if any, at a given distance. We observed how each different particle interacted with the same type of particle, as well as the others, and found that a force is present for five out of the six possible interactions. We found that, while some of the interactions result in an attractive force and some a repulsive force, for all five of the interactions where a force is present, the magnitude of the force is dependent on the inverse square of the distance between them.

Appendix

Distance (m)	Inverse Distance Squared (1/m^2)	Red Red (attract) (N)	Red Green (repel) (N)	Red Blue (neutral) (N)	Green Green (attract) (N)	Green Blue (attract) (N)	Blue Blue (repel) (N)
0.25	16.00	-14.18	14.80	0.00	-17.14	-15.80	29.24
0.50	4.00	-3.88	3.99	0.00	-4.15	-3.84	7.95
1.00	1.00	-1.00	1.00	0.00	-1.00	-1.00	2.00
2.00	0.25	-0.25	0.25	0.00	-0.25	-0.25	0.50

Above is our raw data set for the particle interactions.

Introduction

Our experiment seeks to investigate the properties of "Monke particles," which are particles that exist solely in their own *Pocket of the Manifold*. We aim to identify and characterize interactions that occur between Monke particles, offering qualitative and some rudimentary quantitative observations and data regarding their behaviour. Due to their behavioural characteristics being extremely challenging to test with our given toolset, much of the observations collected are comprised of qualitative observations, with only rough estimates regarding numerical data recording

Methods/Experiments

It is important to note that in the *Pocket* that we were investigating, there was only one specific type of particle, with a certain bound of characteristics which we investigated. We were fairly certain that only one type of particle existed in *Pocket 1976,* as the probability of randomly spawning only one of the two or more potential types of particles to exist is extremely low. In the wave propagation experiment, assuming two particles were to exist and occurred with an equal frequency, the probability of having all 18 particles be of the same type would be $½^{18}$, an order of magnitude of 10^{-6}, which is extremely unlikely.

To gain the most information possible about Monke particles, we decided to conduct a series of experiments to identify their characteristics, each with their own unique experimental methodologies and goals.

The first observation we made with regards to the interactions between Monke particles was that their position vectors relative to one another remained constant under steady state; a system of isolated Monke particles experiencing no external forces had all other Monke particles maintain a constant fixed position relative to one another in space. This also means that if one Monke particle was moved, all other particles would 'follow' it in space.

Moving on from our initial observations, we desired to identify the characteristic of Monke interactions that caused the particles to maintain a fixed position relative to one another. By observing two particles, we were able to conclude that the forces acting on a Monke particle was proportional to its velocity relative to other Monke particles it was interacting with; observations of a moving pair of particles seemed to indicate that as one moved with a nonzero relative velocity to another, they would both accelerate until a zero relative velocity is achieved.

As such, our first hypothesis governing the interaction between two Monke particles was that the force exerted on one Monke particle was proportional to its velocity relative to the other particle.

Relative velocity experiment

In order to test this initial hypothesis, we conducted a test using the equipment provided to determine the relationship between the relative velocity of two Monke particles and the forces they exert on one

another. In this test, the independent variable was the relative velocity of one Monke particle with respect to another, and the dependent variable is the force exerted on the anchored particle.

Relative velocity experiment procedure:

1. Place a Monke particle on a force meter. This will be the 'anchored' particle and will act as the particle whose net force we will be measuring.
2. Hold a second Monke particle in line with the first particle and the force meter.
3. Ensuring that the thumbstick is pushed fully forward, move the Monke particle on the line joining the force meter and the anchored particle.
    a. Observe the range of values the force meter displays; it should vary around a central value. Try to record the central value.
4. Repeat step 3 again to verify that the general range of forces recorded by the force meter is accurate.
5. Repeat steps three and four again, this time depressing the thumbstick fully in the opposite direction to measure the force of two 'arriving' Monke particles.
6. Repeat the whole experimental procedure again, this time depressing the thumbstick slightly, in order to gain qualitative insight with regards to the relationship between the force on a Monke particle and its velocity relative to other particles.
    a. It should be expected that faster travelling Monke particles exert a greater force on other Monke particles.

Observations

What we observed through this experiment was that there was a positive correlation between the relative velocity of Monke particles and the force they exerted on each other; faster moving Monke particles seemed to exert more force. The approximate value obtained for a 'departing' Monke particle at maximum tilt was in the range of 3.1N in the direction of the departing particle's velocity, and the value obtained for an 'arriving' Monke particle was in the range of approximately 2.9N in the direction of the arriving particle's velocity.

The observations we obtained from step 6 were that slower moving Monke particles exerted less force; depressing the thumbstick slightly produced smaller forces as shown on the force meter than if the thumbstick were to be fully depressed. Figure 1 is an image of the experimental set-up of the relative velocity experiment.

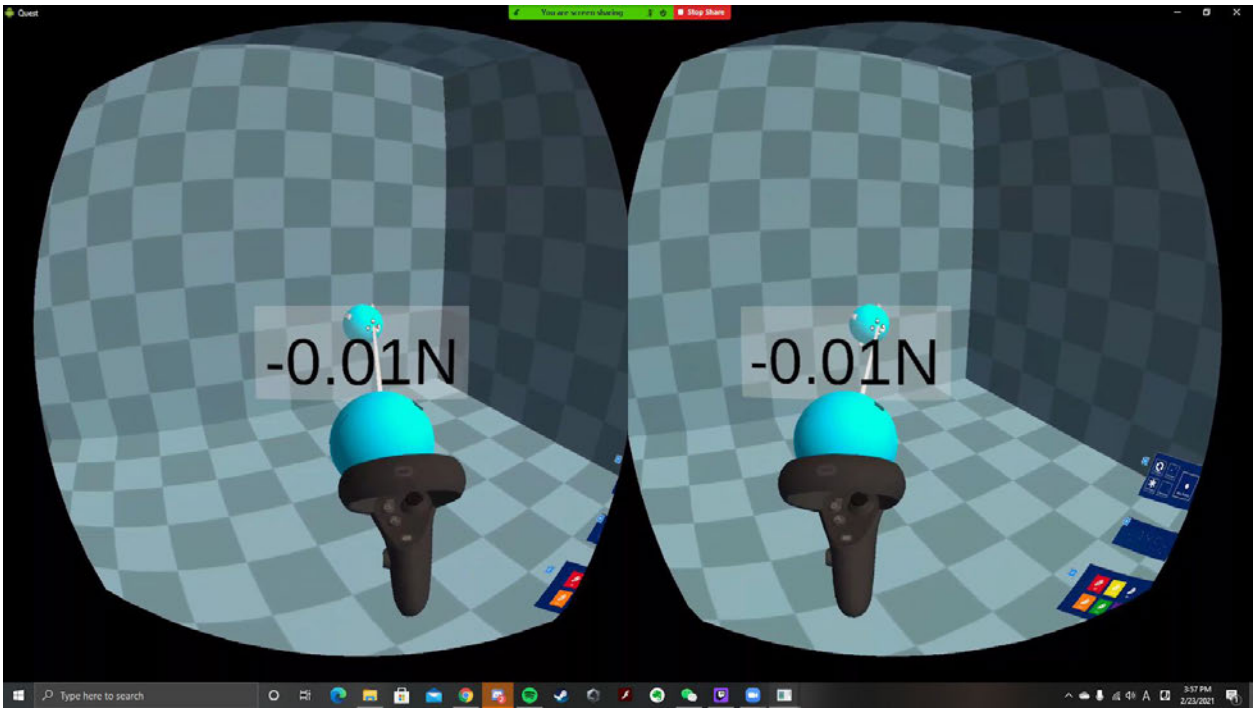


Figure 1: Relative Velocity Experiment

From these observations, we can conclude two things:

1. That force and velocity were positively correlated in Monke interactions, and
2. That the force acting on a Monke particle is in the same direction as other particles' velocities relative to itself.

Nothing can be concluded about the numerical difference between the force on arriving and departing particles; the reading on the force meter varied too frequently to obtain reliable results and therefore the numbers can only represent a 'ballpark' estimate with regards to the value of force exerted.

Conclusion 1: There is a positive correlation between the Monke force and the relative velocity of two Monke particles.

Data analysis

The limitations of this experiment specifically are as follows:

- Bringing the force meter, anchored particle, and particle to be moved proved to be challenging, especially the aspect of ensuring that the force meter's probe was in line with the line joining the two particles
- There was no way to concretely take note of the relative velocity of the two particles, as the necessary equipment to do so was unavailable at the time of the experiment.

In order to build upon these, some of the limitations present must be eliminated. We believe that the most significant limitation to this experiment was the inability to measure the speed of the moving

Monke particles. As such, for future experiments, we encourage researchers to invest in ways to both measure and control the speed of moving Monke particles in order to properly obtain numerical results for their independent variables.

If any experiment conducted in this lab were to be designated the 'most important', this would be it, as it highlights the single defining factor of Monke particles, in that the relationship between their interactions is determined by their relative velocities. All other experiments conducted were following this observation, and thus this takeaway is the most integral when reporting upon Monke interactions.

Slingshot Experiment

In addition to these more concrete experiments that we conducted, we also attempted an additional 'slingshot' experiment to determine if a maximum force between two Monke particles existed. Not many valuable observations were made from this experiment as the equipment used wasn't sufficient, but we think it worthwhile to record the methodology in the event that other researchers have the materials to make observations using this method. Figure 2 is an experimental set-up image of the attempted slingshot methodology. Furthermore, a video of the slingshot is included as a viewable link here.

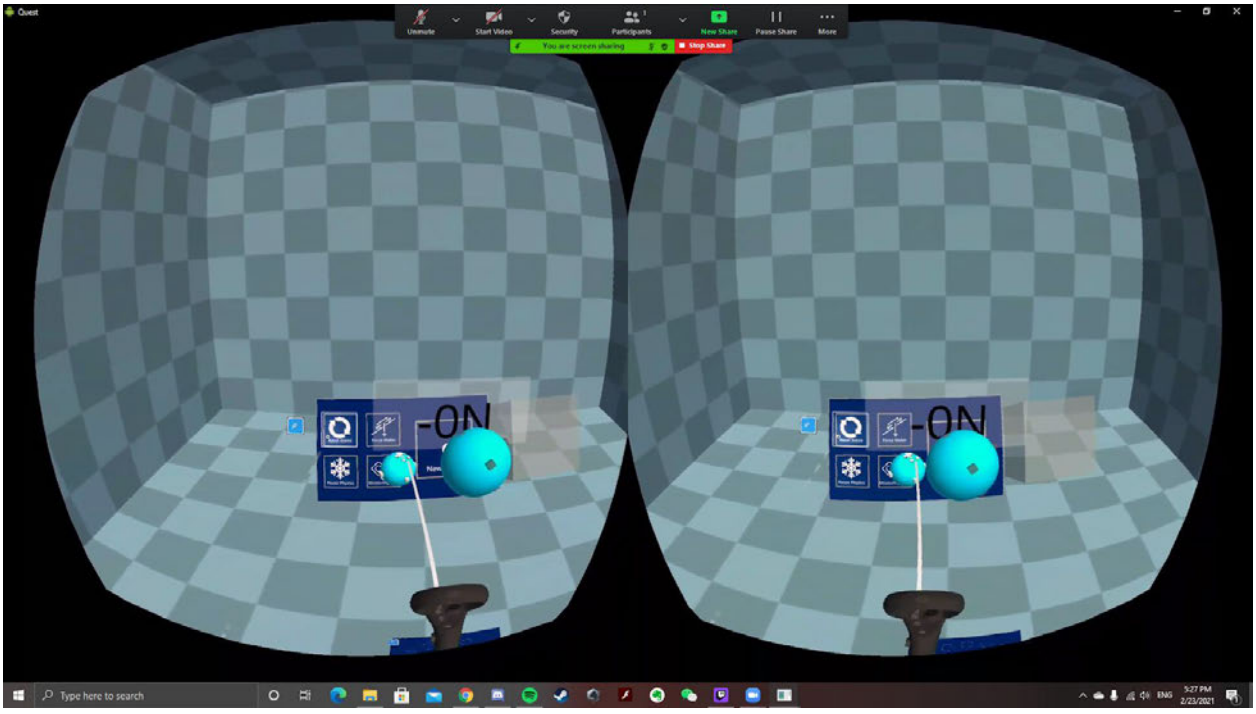


Figure 2: Slingshot Experiment

Slingshot methodology

1. Set up the experiment as described in the procedure description above

2. Instead of using the thumbstick to propel the moving Monke particle at a known velocity, hold it against one of the boards available in the Manifold and use the thumbstick to 'push' the particle into the board
    a. In our attempts, the best board we found was the Spawning board, as the intersection of the gaps between the buttons made a good place to hold the location of the particle steady
3. Pushing the particle into the board 'built up' tension in the connecting line between the particle and the controller, so when the board is removed, the particle is accelerated to a very high velocity.
4. If possible, measure the force of the force meter. If the held particle disappears, there is a high likelihood that it managed to escape the collision detection of the laboratory walls, giving it a lower bound of 7.2m/s velocity.

Separation experiment

We conducted a test in order to test the dependance of the force between Monke particles on their separation. The experiment consisted of measuring and comparing the force between two Monke particles having constant relative velocities and different magnitudes of separation. This test builds upon the relative velocity one by identifying if there were multiple factors that affected the Monke force on particles; it highlighted that the determination of the Monke force was independent of the separation between particles.

Separation experiment procedures:

1. Place a Monke particle on a force meter. This is the 'anchored' particle, which will be fixed throughout the experiment and will be used to measure the force experienced by Monke particles under different separations.
2. Create another Monke particle and hold it at a particular distance from the anchored particle.
3. Now move the new particle in the line of the anchored particle and the force meter with a constant velocity. Record the force experienced by the anchored particle during this motion using the force meter.
4. Observe the force exerted on the anchored particle as a function of time; as the second particle is moving, its separation is changing. Observe the relationship between the time of the movement and the force on the anchored particle
    a. If the force is independent of distance, the proxy indicator of time also shows an unchanging value, so long as relative velocity of the two particles is held constant

Observations:

We observed that the force experienced by the Monke particles was the same in each test as long as the relative velocities were kept constant. The force meter connected to the anchored Monke particle displayed a constant value throughout the movement. Through these observations, we concluded that the force experienced by Monke particles due to each other has no correlation with the separation

between them or that the effect of the separation is so minute that it couldn't be observed using the tools available to us.

Limitations of this experiment were the same as the limitations of the initial relative velocity experiment above; it was challenging to ensure that the force meter and both Monke particles were perfectly in-line, which caused the value of the recording to vary slightly. Furthermore, it was also difficult to ensure that a constant, appropriate velocity was used, as maximum tilt on the controllers gave the particles too high of a velocity which made it impossible to take measurements. Thus, an intermediate velocity was used and attempts were made to ensure that it was kept semi-constant through the use of muscle memory.

As with the relative velocity experiment above, it was again challenging to control the velocities of the particles. Thus, looking into equipment to control and record the velocities of the particles would be integral to obtaining proper numerical results.

Conclusion 2: The Monke force acts with the same magnitude irrespective of distance; the only variable found that changes the Monke force is the velocity of one particle relative to another.

Ruler experiment

The ruler experiment is intended to test the direction of the force exerted from the movement of one Monke particle on the other, and their proportionality of movement. This experiment also builds upon the relative velocity experiment; the first experiment simply determined the magnitude of the force, whereas this experiment aimed to determine its directionality as well. Figures 3 and 4 demonstrate the expected interactions occurring before and after movement in the ruler experiment, as well as the set-up used.

Figure 3: Ruler Experiment Pre-Movement

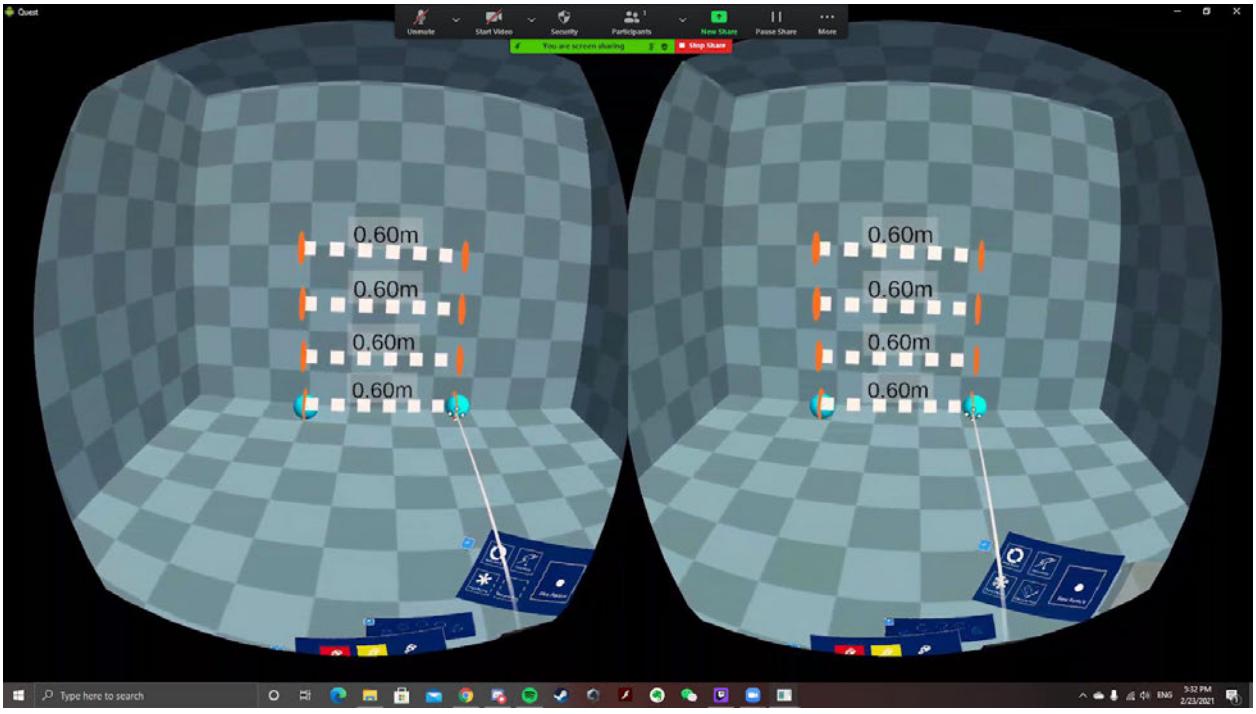


Figure 4: Ruler Experiment Pre-Movement

Ruler experiment procedures:

1. Setup three rulers which are: equal lengths, equidistant from each other, and parallel, so that all rulers lie in a single plane. This is most easily done with the rulers set horizontally, so that observations can be made easily.
2. Place a Monke particle on each end of the topmost ruler.
3. Move one Monke particle from the end of the topmost ruler to the same end of the bottommost ruler.
4. Observe the motion of the second Monke particle.
5. Repeat steps 1-4 for a range of lengths.

Observations

In performing this experiment, we found that the second Monke particle also moved from the end of the topmost ruler to the same end of the bottommost ruler for all lengths attempted. This demonstrates that the force experienced by a Monke particle is in the direction of the velocity of the Monke particle that is exerting the force and does not act on a line joining the two particles. Thus, the force on two Monke particles is not central and instead acts in a direction parallel to the relative velocity of all other Monke particles.

Conclusion 3: The Monke force is a non-central force; the direction of the force on a Monke particle is in the direction of the velocity of all other Monke particles relative to itself.

Data analysis

Looking at the setup used, it is clear that there was relatively large uncertainty with respect to the measurements taken. The rulers used were reporting the same length to two decimal places, however, due to a low level of depth perception, it was impossible to verify that the rulers were fixed in appropriate locations for measurements to be taken well. Furthermore, it was impossible to take accurate measurements of the particles' relative positions as the size of the particles meant that determining the center of each particle could only have been done by eye, and not to a high degree of precision as human judgement for the 'center' of each particle could vary from experimenter to experimenter.

Furthermore, there may have been a minute central force that was being exerted on the two particles. However, as a result of the short distance between the topmost and bottommost rulers, the magnitude of these forces may have been imperceptibly small. Hence, should future experimenters wish to conduct research on this phenomenon, we would encourage them to create an experimental method that enables them to move the particles over a longer distance to test this hypothesis to a better degree of accuracy. Furthermore, variations in the separation between particles may have affected the centrality of the force; experiments into how the separation factors in would also be useful knowledge to obtain.

Superposition experiments

In order to test the possibility of force superposition in the Monke interactions, we conducted two experiments to determine the superposition of multiple Monke forces. Each experiment demonstrated

both destructive superposition, where the forces by multiple Monke particles on one particle cancelled out, and constructive superposition, where the forces by multiple Monke particles resulted in a greater net force on a single Monke particle. This experiment was a fairly standalone experiment; it simply aimed to determine if the Monke force acted as other non-contact forces did outside the *Manifold*, in that they superpose. Figures 5 and 6 demonstrate the two set-ups used for the superposition experiments, as well as an included video of one of them here.

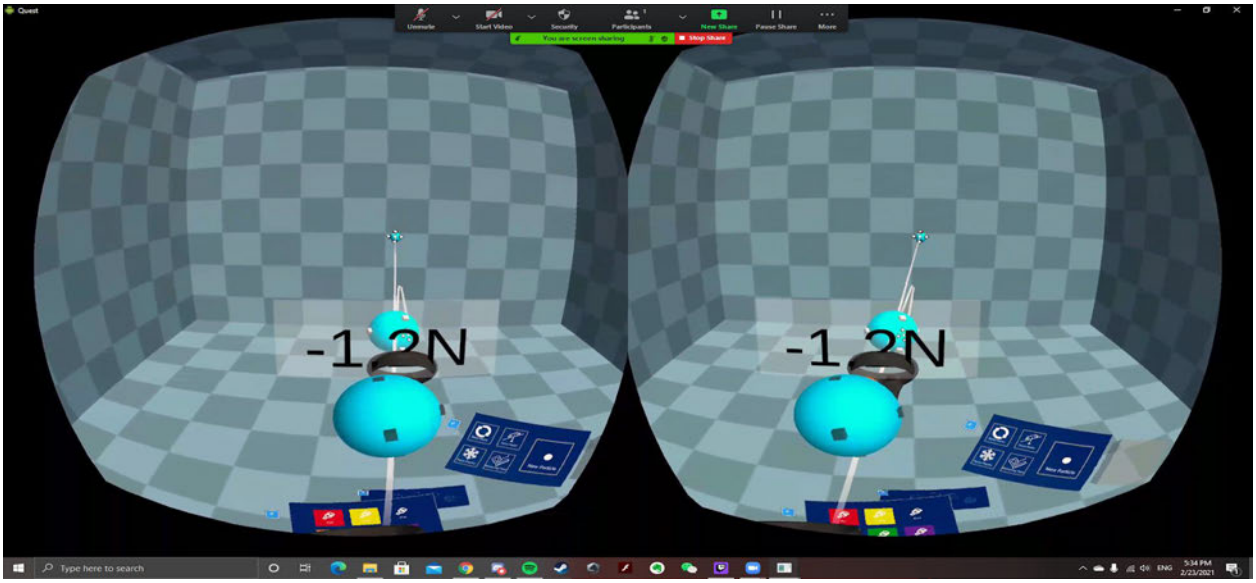


Figure 5: Superposition Experiment 1

Figure 6: Superposition Experiment 2

Superposition experiment procedures:

1. Place a Monke particle on a force meter. This is the 'anchored' Monke particle, which will be used to measure the force experienced by Monke particles under different circumstances.
2. Create another Monke particle some distance away from the anchor and hold it with a controller, close to the user.
3. Push the held Monke particle away from the particle at constant speed by pushing the controller joystick as far up as it will go. Ensure that the particle will not collide with the anchor.
4. Record the force experienced by the anchored Monke particle from the force meter during this motion.
5. Create another Monke particle; hold each with a controller, close to the user.
6. Simultaneously push both held Monke particles away from the particle at constant speed by pushing the controller joystick as far up as it will go. The lines traveled by the Monke particles should be parallel, in order to observe the presence of superposition.
7. Record the force experienced by the anchored Monke particle from the force meter during this motion.
8. Repeat steps 3-7 until it is evident that the magnitude of the force associated with the movement of two Monke particles is either the same or distinctly different from the force associated with the movement of one Monke particle.
9. Position the two held Monke particles so that one is close to the user, and one is very far from the user. The position vectors made by the user and the particles should be parallel.
10. Simultaneously push the close Monke particle away from the user at constant speed, and pull the far Monke particle toward the user at constant speed.
11. Record the force experienced by the anchored Monke particle from the force meter during this motion.

Experiment 2:

Using the anchored particle and two held particles from the previous experiment:

1. Hold the two Monke particles some distance away from the anchor.
2. Simultaneously move the held particles in opposite continuous vertical oscillatory motion: each particle should be continuously moved up and down so that the particles are always moving in opposite directions. The paths traveled by the Monke particles should be parallel.
3. Record the force experienced by the anchored Monke particle from the force meter during this motion.
4. Repeat steps 2-3 until it is evident that the magnitude of the force associated with the movement of two Monke particles is either the same or distinctly different than the force associated from the movement of one Monke particle.

While it is not necessary to perform both experiments in order to prove the presence of superposition, performing both experiments provides a more intuitive understanding of the presence of superposition within the motion of Monke particles.

Observations and data analysis

By performing these experiments, we came to the conclusion that the force experienced by Monke particles, and thus the motion of Monke particles, follows the law of superposition. We found that when two Monke particles were moved in the same direction, the force experienced by the anchor was significantly greater than the force experienced by the anchor when one Monke particle was moved. When two Monke particles were moved in opposite directions, the force experienced by the anchor was close to zero, which is the expected value under superposition for two particles moving with equal magnitudes and opposite directions. The presence of superposition also demonstrates that within a system that contains multiple Monke particles moving in different directions, the force experienced by a separate particle is in the direction of motion of the center of mass of the system.

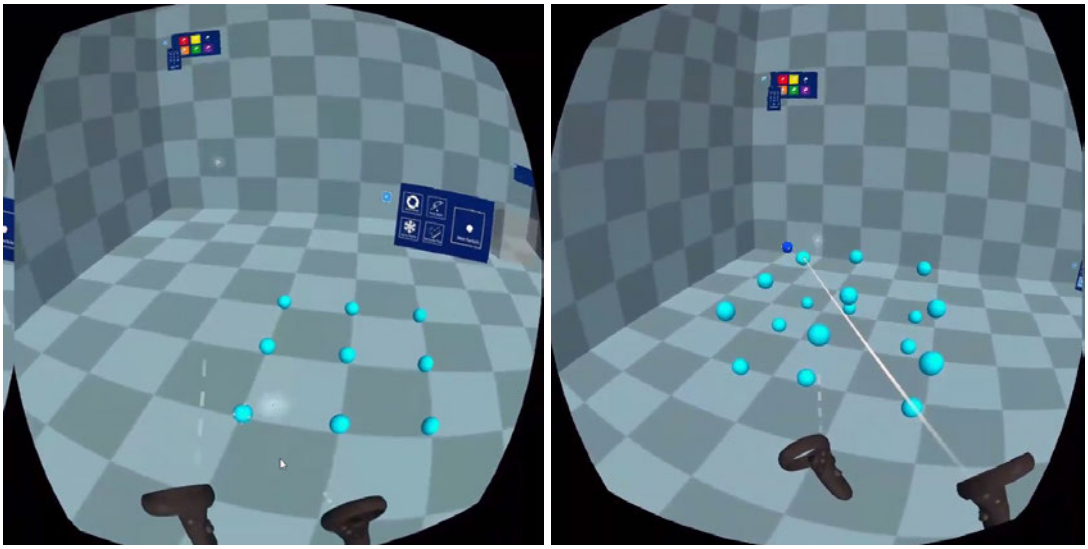
It was difficult to determine if the forces added in the same way that electrostatic or gravitational forces act, as it was a challenge to make two Monke particles travel at the exact same speed. As the initial recordings between a single pair of particles was already relatively unreliable numerically, these recordings were doubly numerically unreliable, depending on the relationship between two pairs of interactions. However, it can be qualitatively concluded that the principle of superposition is true for Monke interactions, as the forces observed were outside that of the expected range with only a single particle.

Conclusion 4: The Monke force acts with field superposition; forces additively and destructively superpose.

Wave propagation experiment

One of our most conclusive experiments was the wave propagation experiment that we conducted to identify any potential propagation delay, or the wave effect, in the forces between Monke particles.

Figures 7 and 8 illustrate the experimental set-ups used for the wave propagation experiments.



Figures 7 and 8: Wave Propagation Experiments in Two and Three Dimensions

Wave propagations experiment procedures:

Notes:

Monke square - a 3x3 square at the bottom of the simulation made by aligning Monke particles on each corner of a square in the simulation.

Monke cube - a 3D-cube made by making a cube in one of the corners of the simulation.

Lead Monke particles - the Monke particle being moved by the VR investigator, using the controller.

Following Monke particles - the Monke particles experiencing motion due to the movement of the lead Monke particle.

1. Knowing that the simulation scenario has a grid-like background, align the Monke particles each on the corners of the grid squares in frozen physics mode in order to limit variation in movement during placement of the Monke particles. This should be done until a 3x3 of Monke particles is created, making it cover 4 squares, 2 squares horizontally, and 2 squares vertically.
2. Unfreeze physics, and slowly move a lead Monke particle from the corner of the grid square to another grid square.
3. It is important to note and make observations on the position of the following Monke particles which are not being displaced directly by the VR investigator.
4. Try moving a different Monke particle, which is not the original lead Monke particle from the one grid corner to the other, the Monke square should follow.
5. Assign a lead Monke particle which is closest to your favorite corner of the simulation into your favorite corner. The following Monke square should align perfectly into the square as they did in previous steps.
6. Make a cube with the Monke particles, by aligning the Monke particles in a square pattern, approximately a grid-square height above the original Monke square.
7. Assign a Monke particle to be the lead which is at the base of the Monke cube with respect to the vertical. Move the lead Monke particle to a square close to you, the cube should follow.
8. Additionally, it may be fun to move the lead Monke particle vertically as well.

Observations

This experiment was designed to test whether there may be some wave propagations interactions exhibited on the particles. Although, this was not true, and it also showed that, moving a lead Monke particle which is a set distance from other Monke particles will allow that Monke particle to maintain that specific distance from the other following Monke particles without delay. This experiment is to analyze a complex hypothesis which is built upon other hypotheses that were developed originally in

this experiment. The Monke particles do not act with wave propagation with respect to each other, but act in a delay.

Conclusion 5: Monke particles do not act on a distance-based delay. By eye, the time it takes for the interaction to occur between a pair of Monke particles is the same irrespective of distance.

Data analysis

There may have been an error introduced into the data which is due to not enough Monkey particles being included into the experiment. With not enough Monke particles, a faster wave propagations delay will not be visible in this case, since a Monke cube has minimal delay. Furthermore, as no numerical method to obtain time data was present, the majority of these measurements were only true to the extent that they can be observed by the human eye, as there may be an extremely small delay in reality.

Interpretations & Conclusions

Overall, the experiments conducted provide comprehensive qualitative data with regards to the behaviour and interactions of Monke particles. The observations are quite expansive, spanning interactions between simple pairs of Monke particles to complex systems of particles as well. We were able to set down several key relationships of Monke particles; unfortunately we're unable to collect numerical data as a result of the limited availability of data measurement equipment we were provided, paired with the difficult-to-measure qualities of the Monke particles.

As such, all the data that was obtained was exclusively qualitative in nature; no concrete relationships nor mathematical models were derived as a result of the lack of numerical data obtained, though it is speculated that the force experienced by a Monke particle is directly proportional to the relative velocity between it and the Monke particle exerting the force. Thus, future experimentation that takes place must be conducted with the goal of quantitative verification of the observations we have made in our research. For this, we encourage future researchers to obtain equipment that enables them to determine the velocity of objects moving in the *Pocket of the Manifold* in order to develop well-set dependent and independent variables for quantitative experimentation.

Furthermore, for the vast majority of the experiments conducted, many of the readings we took were unable to be numerically verified, as the appropriate measuring equipment was unavailable for use. Thus, much of the conclusions that we drew was 'by eye'; nothing numerical can be said about them and their validity is to the extent that they are detectable by the human eye and no further.

List of conclusions drawn:

1. The force acting on two Monke particles is positively correlated with their relative velocities: faster-moving particles exert forces greater in magnitude.
2. The Monke force is independent of Monke particle separation; two particles interact the same way regardless of the distance between them.
3. The Monke force is not a central force; the direction of the force exerted on one particle is determined only by the velocity of other moving Monke particles relative to itself.

4. The Monke force is cumulative; superposition of the force does exist and works in a summative interaction.
5. The Monke force has no time propagation delay; by eye, all Monke particles experience the same force at the same time in the presence of another moving particle, irrespective of distance between them.

Finally, we felt as though a note regarding the naming of the Monke particle should be included as an endnote to this report. Based on our initial observations, we determined that the particles followed each other, and off of this, we determined that the 'Monke particle' would be an appropriate name for such a particle, as "monkey see, monkey do".

Appendix B

EXAMPLE LAB MANUALS

Lab A2 - Model Testing: Coulomb's Law

Last week, you qualitatively tested Coulomb's Law using simple equipment. Your goal in this lab today is to quantitatively test whether the interactions between charged particles in a virtual reality (VR) lab obey Coulomb's law.

Note: Physics in the virtual lab follows a limited set of models. Most of these models are based on the real world, but none are exactly accurate. It's impossible to perfectly simulate real-world physics, and many rules have been deliberately changed for the next two labs.

Learning Objectives

1. Design and conduct an experiment with well determined independent and dependent variables to test if the simulated charged particles interact according to Coulomb's law.
2. Estimate random or instrumental uncertainties for measurements.
3. Create a set of graphs to understand how to linearize a graph.
4. Make a judgment of the validity of the model being tested based on a well-reasoned argument.

1. Explore the Simulation

Make sure to take turns using the VR headset so that all team members who want to try can use it.

1.1 VR headset safety

- The headsets should be used only while seated.
- Students in a headset MUST be given adequate space to move their body and swing their arms. They cannot see their fellow students. Headset users ALWAYS have the right-of-way.
- Always ask for consent before touching anyone in a VR headset (e.g. to put their finger on a controller button).
- Wipe down the headset (with the wipes provided) between users, every time the headset changes hands.
- No one is required to use the headset. If a student experiences nausea or other discomfort while using the headset, they should take it off immediately and notify an instructor.

1.2 Oculus VR Introduction Guide – Optional

If you are new to using the Oculus VR system, an introduction video and document, *Oculus VR Introduction*, are available on Canvas under "Lab A2 – in lab".

1.3 Wifi – off

Please keep the Wifi Off when using the Oculus VR Headsets, even if the software prompts you to turn it on. If the Wifi is turned on it will disable the lab NOMR software. You may see a notification on screen regarding "Multiple Licenses" for the NOMR App. If you see this, then the Wifi has been turned on. To turn the Wifi off, go to the left side of the tool bar at the bottom of the screen. Click on the Battery/Wifi logo. Then select "Wifi" and toggle the slider to "Off". Ask your instructor if you need assistance.

1.4 Launch the Simulation (VR headset operators)

1. Make sure the Oculus headset is connected to the Desktop PC with the braided USB cable.
2. Power on the Oculus headset by pressing and holding the button on the right side of the headset.
3. You will see a message "Allow access to data". Select "Allow" using either hand controller.
4. Click "Create Guardian" and follow directions to set ground level. (If prompted to, select "Stationary Boundary" and confirm).
5. From the Apps menu open the NOMR app (it should be at the top left of the screen).
6. Ask a lab-mate to double-click to open the *Quest_Mirror-Shortcut* on **the desktop computer screen**.

You should now be in the virtual lab, and the screen should be mirrored on the Desktop PC. It is sometimes necessary to re-open the *Quest_Mirror* software on the desktop. If the *Quest_Mirror* software is not functioning, check to be sure the USB cable is connected and repeat from step 3 above. If you have any difficulties, your lab instructors should be able to help you.

1.5 Familiarize yourself with the VR lab environment by completing the following steps:

1. Enter the code "1856" into the keypad and press Submit to load today's lab setup.
    a. Press buttons by pointing at them and clicking the trigger button on your controller.
2. Create a new positive charge by pointing at the button labeled "Positive" and clicking the trigger. Note: each particle has the same amount of charge, $\pm q$.
3. Point at the newly created charge and hold down the grip button to hold the charge and move it around your space.
4. While holding the charge, push the controller thumbstick away from you to push the charge away. Move the thumbstick toward you to pull the charge toward you. This will work for every movable object in the simulation.
5. Anchor the charge by pointing at it and pressing the trigger. This locks the charge at a point in space and is visually represented by a cage around the charge. To un-anchor the charge, point at it and press the trigger again.
6. Pause physics in the simulation by clicking the button labeled "Pause Physics." This pauses the motion of all free charges, allowing you to configure an experiment. Clicking the same button again will resume physics.
7. Delete the charge by holding down the B or Y button, pointing at it, and pressing the trigger.
8. Reset the lab, deleting all charges and tools, by clicking the "Reset" button.

1.6 Electrically Charged Particles

Take turns exploring the simulation. Try creating multiple electric charges. Try mixing charge types. Try adding more than two charges to the simulation. Discuss with your team how these simulated electric charges interact with each other. Does Coulomb's law align with your observations of the simulated charges?

2. Testing Coulomb's Law

The goal for the end of today's lab is to test a model describing the force between two electrically charged particles in the VR lab. To do so, you will design an experiment to answer the following question: do your observations of the simulated charged particles align with Coulomb's law: $F = kq_1q_2/r^2$?

2.1 Tools Available for Experimental Design

Before testing a model, we must first understand the tools available to use in experiments.

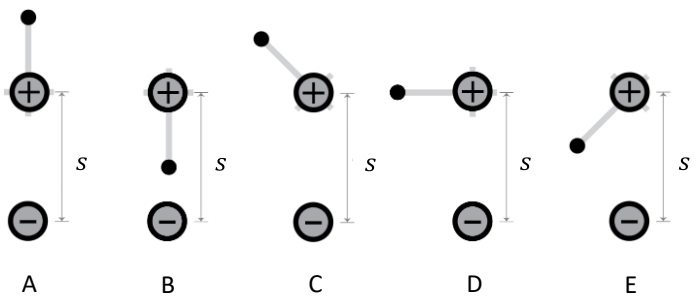
Create a **measuring tape**. Create two charges and anchor them nearby. Measure the distance between the charges. Discuss the following questions with your team:

- What is the uncertainty associated with your measurements?
- Is this random uncertainty or instrumental uncertainty?

Create a **force meter**. Attach a positive charge to the force meter by placing the charge on the spiky end of the meter (note: the orange end of the force meter is always anchored). Anchor a negative charge near the force meter. If nothing happens, you may need to resume physics in the simulation.

The force meter acts like a spring force meter where the spring can be stretched or compressed (if you are unfamiliar with a spring force meter ask your instructor to show one to you). The force meter will pivot around the orange, anchored, end of the force meter until equilibrium is reached. This can take a few seconds.

In the diagrams at right, a positively charged particle is attached to a force meter with various orientations, and a negatively charged particle is anchored a distance s away from the positive charge. Note that as the separation of two charges is the same in all cases, the electric force on the positively charged particle due to the negatively charged particle must be the same in all cases and point downward as shown.



Discuss with your team and experiment with the force meter in VR to answer the following questions in your lab notebook (Please note, except for diagram A, these configurations are not in equilibrium. So, you will need to take approximate readings as the particle moves and reaches equilibrium.):

1. Rank the diagrams in order of the value displayed by the force meter at the instant the diagram represents (with +/- sign); if any quantity is zero, state so explicitly.
2. Does a positive reading on the force meter always represent an attractive force, always represent a repulsive force, or can it represent either?
3. Can you relate the approximate force meter's reading with the electric force vector and the direction of the force meter shown in the diagrams above?
4. How are the magnitude and direction of a force indicated by the force meter?
5. What is the uncertainty associated with your measurements? Is this random uncertainty or instrumental uncertainty?

Try step 1 above with two positive charges discuss if your measurements are consistent with your discussion for steps 2, 3, and 4 above.

CHECK-IN 1: ***Discuss with an instructor your understanding of the available measurement tools before proceeding into your experimental design.***

2.2 Designing an Experiment

Now, discuss with your team how you will experimentally observe the interactions you are interested in. Brainstorm the potential experimental variables that you could adjust in electric charge interactions in the VR. Make a list of these variables and assign each variable a symbolic representation to use as shorthand. Indicate whether the variable can be independent, dependent, or controlled. In general, the uncertainties for the measurements of the independent variables and control variables should be much smaller than those for dependent variables. The table below can be used as a template for your list.

Variable	Symbol	Can it be independent?	Can it be dependent?	Can it be a control?
Separation	r	Yes	No	Yes
Force	F			

Consider the following scientific questions:

- How does the force depend on the charge of the particles?
- How does the force depend on the separation of the particles?

Discuss with your team how you would collect data to answer each of these questions. For each experiment, what is the independent variable and the dependent variable? Are there any control variables? How do you estimate uncertainties of your measurements?

- Do you expect the force vs. charge graph to have a linear relationship based on Coulomb's law?
- Do you expect the force vs. separation graph to have a linear relationship based on Coulomb's law?

CHECK-IN 2: ***Discuss your experimental design with an instructor before proceeding.***

To learn special data analysis techniques for exploring non-linear relationships, let's test Coulomb's law by conducting an experiment to answer the second question. After discussion with the instructor, proceed to make measurements.

3. Data Analysis

3.1 Graph Linearization

In physics, we often investigate how one quantity relates to another by graphing data. If your data graphs as a curve, the quantities have a non-linear mathematical relationship (e.g. $y = x^2$). Non-linear data is more complicated than linear data to analyze. It is hard for the human mind to accurately recognize any specific function other than a straight line. We can take advantage of the human capacity to recognize straight lines by transforming the data such that what we graph on our axes results in a linear relationship.

Your goal is to create a linearized graph to see if your data agrees with Coulomb's law. Draw rough graphs (no need to include uncertainties for now) of the data you collected for the relationship between the forces, F, and the separation, r using the following axes:

- F (on the vertical axis) vs. r (on the horizontal axis)
- F vs. $1/r$
- F vs. $1/r^2$
- F vs. $1/r^3$

The following website describes how to make a graph on Excel. Scatter Plot in Excel (Easy Tutorial) (excel-easy.com)

Discuss with your team which of the graphs you created looks most linear and why. Compare Coulomb's law, $F = kq_1q_2/r^2$, to a general formula for a straight line, $y = Ax + B$. Which quantity in Coulomb's law corresponds to which quantity in the straight-line formula for the graph that looks the most linear?

Straight line formula	Coulomb's Law
y	F
x	
A	
B	

CHECK-IN 3: ***Discuss these graphs and the table above with the instructor.***

3.2 Graph your data

Create a complete graph with uncertainty bars for the most linear relationship using a graphing tool like Google Sheets or Excel. **You should choose which graphed quantity has the larger uncertainty and include it in the graph.** The following websites have instructions on how to add uncertainty bars (error bars).

- https://excelquick.com/excel-charts/excel-chart-error-bars/ (for Excel)
- https://www.youtube.com/watch?v=8LuKPw0_pMI&ab_channel=BenPaulson (for Google Sheets)

Include a trendline (best-fit line) for your linearized graph. Does the best-fit line go through at least 2/3 of the uncertainty bars? Do the simulated charged particles obey Coulomb's law?

BEFORE YOU GO:

- ***Ask TA to sign off.***
- ***Be sure to turn off the Oculus headset by holding the power button (on the right side of the headset) before putting them away in the box.***

3.3 Optional Activity: Consider the following questions and how to experimentally answer them.

- Do the forces between particles obey superposition?
- What stable structures can you make with these simulated charges?

- If you create a large number of one type of charged particles, how do they distribute themselves in the room? What can this tell us about the distribution of charges on conductors with sharp edges?

DUE this Friday

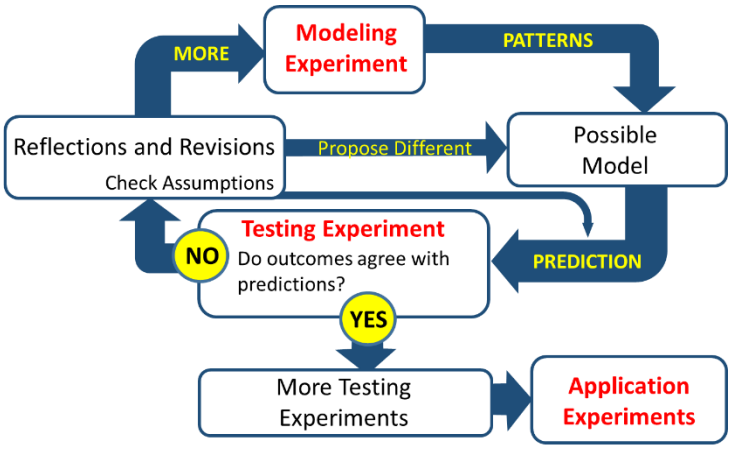
Do Lab HW 2 on Canvas before 11:59 PM on Friday. This assignment will be graded individually.

DUE next Friday

A full lab report should have introductions, methods, data analysis, and discussion/conclusions sections as discussed during the lab workshop in the first week of the quarter. You can find a sample lab report on Canvas under Files/Lab handouts/Sample Lab Report (Empirical Modeling).pdf. Before 11:59 PM next Friday, you'll write with your team members and submit **only the methods section** of a lab report for today's experiment. Note that you do not need to write other parts of a report. See the rubric under **Lab A2 – methods** on Canvas.

Lab A3 - Empirical Modeling: The Minty Particle

In the last two labs, you conducted model testing experiments; you tested Coulomb's Law. Your goal in this lab is to perform another kind of experiment, an empirical **modeling experiment**, to create a model describing the interactions between never-before-seen fictitious particles (dubbed "Minty Particles") in a virtual reality (VR) lab that do not exist in the real-world.



Learning Objectives

1. Make qualitative observations of general behavior of the fictitious "Minty particles" that are simulated in the VR.
2. Design and conduct an experiment with well determined independent and dependent variables.
3. Estimate random or instrumental uncertainties for measurements.
4. Create a linearized plot.
5. Synthesize the observations and measurements to develop an empirical model that describes the "minty" particle interactions.

1. Team Reflection

Before starting today's lab, each member of the team should share what is working well in the team. And if any of you feel that something is not working well, for instance, if someone is not behaving constructively, or if someone is not doing their fair share of work, share this with the team and discuss possible solutions.

2. Explore the Simulation

2.1 VR headset safety

- The headsets should be used only while seated.
- Students in a headset MUST be given adequate space to move their body and swing their arms. They cannot see their fellow students. Headset users ALWAYS have the right-of-way.
- Always ask for consent before touching anyone in a VR headset (e.g. to put their finger on a controller button).

- Wipe down the headset (with the wipes provided) between users, every time the headset changes hands.
- No one is required to use the headset. If a student experiences nausea or other discomfort while using the headset, they should take it off immediately and notify an instructor.

2.2 Oculus VR Introduction Guide – Optional

If you are new to using the Oculus VR system, an introduction video and document, *Oculus VR Introduction*, are available on Canvas under "Lab A3 – in lab".

2.3 Wifi – off

Please keep the Wifi Off when using the Oculus VR Headsets, even if the software prompts you to turn it on. If the Wifi is turned on it will disable the lab NOMR software. You may see a notification on screen regarding "Multiple Licenses" for the NOMR App. If you see this, then the Wifi has been turned on. To turn the Wifi off, go to the left side of the tool bar at the bottom of the screen. Click on the Battery/Wifi logo. Then select "Wifi" and toggle the slider to "Off". Ask your instructor if you need assistance.

2.4 Launch the Simulation (VR headset operators)

1. Make sure the Oculus headset is connected to the Desktop PC with the braided USB cable.
2. Power on the Oculus headset by pressing and holding the button on the right side of the headset.
3. You will see a message "Allow access to data". Select "Allow" using either hand controller.
4. Click "Create Guardian" and follow directions to set ground level. (If prompted to, select "Stationary Boundary" and confirm.)
5. From the Apps menu open the NOMR app (it should be at the top left of the screen).
6. Ask a lab-mate to Double-click the *Quest_Mirror-Shortcut* on the desktop screen.

It is sometimes necessary to re-open the *Quest_Mirror* software on the desktop. If the *Quest_Mirror* software is not functioning, check to be sure the USB cable is connected and repeat from step 3 above.

You should now be in the virtual lab, and the screen should be mirrored on the Desktop PC. Enter the code "**1550**" into the keypad and press Submit to load today's lab setup.

2.5 Familiarize yourself with the new VR lab environment

In a simulation, we are not bound by the constraints of working with real-world physics. A new fictitious particle, the "Minty Particle", has been invented and programmed into the simulation. <u>The characteristics of all the Minty particles programmed in the VR are identical.</u>

3. Minty Particle Interactions

3.1 Qualitative Observation of a New Particle

We will explore minty particles in depth with quantitative tools to develop a model for the forces they exert upon each other. But first, focus on making and recording qualitative observations.

- How do two minty particles interact with each other? Three? More?
- How do minty particles behave at extremes? Short distance? Long distance? How does a system of many minty particles interact with each other?
- What other features of minty particles' behavior do you find to be significant?

CHECK-IN 1: ***Discuss your qualitative observation with an instructor. Obtain a new code to be used for the next part of the experiment from the instructor before proceeding.***

3.2 Design Your Experiment

Your aim now is to develop a quantitative empirical model of the interactions between minty particles using data from an experiment. To design an experiment, you can design-by-analogy using the experiment you conducted last week as a template for modeling an interaction between particles. You will have access to the force and distance measurement tools you used previously with conventional electric charges, and you can create as many minty particles as you need.

Identify the independent, dependent, and controlled variables in your experiment and define symbol for each variable. The table below shows an example of this identification as applied to the charged particles experiment to help guide your thinking about minty particles.

	Charged Particles (example)	**Minty Particles**
Dependent Variables	F	
Independent Variables	Sign of q_2, r	
Control Quantities	q_1, $\lvert q_2 \rvert$, $\hat{r}$	

Your experimental design should also answer the following questions:

- How many data points do you plan to take and for what values of your independent variables?
- How do you distinguish between attractive force and repulsive force measurements?
- What sources of random & instrumental uncertainty are present in your experiment?

CHECK-IN 2: ***Discuss your experimental design with an instructor and then conduct your experiment!***

4. Develop an Empirical Model

Using the data that you gathered from your experiment, your goal is to develop a model for minty particle interactions. Use the table below as a guide to develop your model (the Coulombic model for electric charge interactions is described as an example).

	Charged Particles (example)	**Minty Particles**
Symbolic Representation	$\vec{F} = \frac{kq_1q_2}{r^2}\hat{r}$	
Verbal Representation Define all relevant quantities	• The force between two particles with charges q_1 and q_2 is proportional to each charge and has a constant k. • The force falls off with the square of the distance between the charges r. • The force acts in the $\hat{r}$ direction, along the line connecting the two charged particles. Note: $\hat{r}$ is a unit vector that only indicates the direction.	

As you develop your model, keep in mind:

- You are conducting an empirical modeling experiment. Unlike the model testing experiment that you conducted for Coulomb's law, there is no known model that you are testing. The only "right" answer is the one you make the best case for.
- It's normal practice in physics experimentation to revise your experiment and collect additional data as you learn things and make decisions.
- An equation symbolizes information but does not necessarily capture everything about the model (such as assumptions that you might be making or limitations of the model). Describe any assumptions you are making and the limits over which you believe the equation is valid. Can you, for example, describe the interactions between two minty particles that are infinitely apart?
- One experiment can't fully flesh-out a model. The best model for the minty particle interactions is the one that your team feels best describes them, out of all your team's ideas. You're being assessed based on the quality of your collaborative process - you're not in competition with other teams for the most sophisticated model.
- Refer to last week's lab for how to linearize a plot.

CHECK-IN 3: ***Show the data you collected to an instructor and discuss your model with them before you leave.***

5. Preparing the assignment

Before 11:59PM on Friday next week, you'll write a partial report with a linearized graph (using a plotting tool like Google Sheets or Excel) and conclusion section as a team for your empirical model of the interactions between minty particles. Note that this is not a full report. Also note that you used Coulomb's law as a guide to develop your model, but your partial report should be on your model on the Minty particles, not on Coulomb's law.

Before leaving the lab today, discuss the following:

- When should you meet to prepare the partial report? Recall that we strongly discourage you from dividing up the tasks without discussing the final product.
- Who should do what tasks before the meeting?

BEFORE YOU GO:

- ***Ask TA to sign off.***
- ***Be sure to turn off the Oculus headset by holding the power button (on the right side of the headset) before putting them away in the box.***

Lab A4 - Empirical Modelling: The Exotic Particles

Your goal in this lab is to perform an empirical modeling experiment to develop a mathematical model describing the complex interactions between fictitious never-before-seen particles in a virtual reality (VR) lab.

Learning Objectives

1. Make qualitative observations of general behavior of the fictitious "exotic" particles simulated in the VR environment.
2. Design and conduct an experiment with well determined independent and dependent variables.
3. Estimate random or instrumental uncertainties for measurements.
4. Synthesize the observations and measurements to develop an empirical model that describes the "exotic" particle interactions.

1. Explore the Simulation

1.1 VR headset safety

- The headsets should be used only while seated.
- Students in a headset MUST be given adequate space to move their body and swing their arms. They cannot see their fellow students. Headset users ALWAYS have the right-of-way.
- Always ask for consent before touching anyone in a VR headset (e.g. to put their finger on a controller button).
- Wipe down the headset (with the wipes provided) between users, every time the headset changes hands.
- No one is required to use the headset. If a student experiences nausea or other discomfort while using the headset, they should take it off immediately and notify an instructor.

1.2 Oculus VR Introduction Guide – Optional

If you are new to using the Oculus VR system, an introduction video and document, *Oculus VR Introduction*, are available on Canvas under "Lab A4 – in lab".

1.3 Wifi – off

Please keep the Wifi Off when using the Oculus VR Headsets, even if the software prompts you to turn it on. If the Wifi is turned on, it will disable the lab NOMR software. You may see a notification on screen regarding "Multiple Licenses" for the NOMR App. If you see this, then the Wifi has been turned on. To turn the Wifi off, go the left side of the tool bar at the bottom of the screen. Click on the Battery/Wifi logo. Then select "Wifi" and toggle the slider to "Off". Ask your instructor if you need assistance.

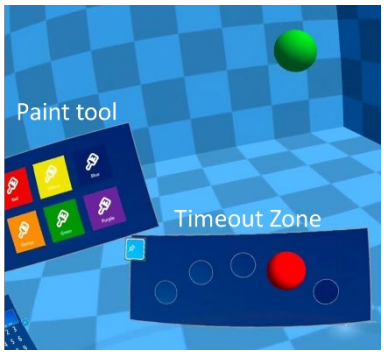
1.4 Launch the Simulation (VR headset operators)

1. Make sure the Oculus headset is connected to the Desktop PC with the braided USB cable.
2. Power on the Oculus headset by pressing and holding the button on the right side of the headset.
3. You will see a message "Allow access to data". Select "Allow" using either hand controller.
4. Click "Create Guardian" and follow directions to set ground level. (If prompted to, select "Stationary Boundary" and confirm.)
5. From the Apps menu open the NOMR app (it should be at the top left of the screen).
6. Have a lab-mate double click the *Quest Mirror-Shortcut* on the desktop.

It is sometimes necessary to re-open the *Quest_Mirror* software on the desktop. If the *Quest_Mirror* software is not functioning, check to be sure the USB cable is connected and repeat from step 3 above.

You should now be in the virtual lab, and the screen should be mirrored on the Desktop PC. **You will need two codes today to enter in the VR lab.** First, enter the code "2350" into the keypad and press "Submit" to load today's lab setup. For the second code, ask your TA for a 4-digit code specific to your team.

1.5 Familiarize yourself with the new VR lab environment



Your group's VR lab comes with one to three distinct kinds of particles, which are generated at random each time the "New Particle" button is pressed. All particles will look the same on creation. You can use the "Paint" tool to color them. Coloring them does not change the nature of the particles. The lab features a Timeout Zone with 5 spots to place particles; while in timeout, those particles will neither exert forces on other particles nor experience forces themselves.

2. Exotic Particle Interactions

2.1 Qualitative Observations

Conduct qualitative experiments to answer the following questions:

- How many unique types of particles does the VR lab generate?
- How can the unique physical behaviors of an unidentified particle be determined? Recall from Lab A2 that you explored two types of charge and deduced that the same types repelled each other while the different types attracted each other. In today's lab, your particles are not strictly positive or negative charges, or follow the rules for charged particles: "like charges repel; opposite charges attract."

Document the setup and results of each experiment. Come up with names to reference each particle.

CHECK-IN 1: DISCUSS YOUR FINDINGS WITH AN INSTRUCTOR.

2.2 Design Your Experiment

Once you have identified your particle(s), develop and conduct an empirical modeling experiment to characterize their interactions. Here is a non-exhaustive list of questions to get you started:

- What independent variables affect the force between a pair of particles?
- What is the functional relationship of each independent variable to the force between particles?

CHECK-IN 2: ONCE YOU DESIGN YOUR EXPERIMENT, DISCUSS WITH AN INSTRUCTOR.

2.3 Build Your Model

Develop a mathematical model for the interactions between particles that is consistent with your answers to the above questions and supported by the qualitative and/or quantitative experimental evidence. Note that your model can be composed of results from more than one empirical modeling experiment, so long as those results are consistent with one another.

CHECK-IN 3: SHOW AN INSTRUCTOR YOUR DATA AND DISCUSS YOUR MODEL WITH THEM BEFORE LEAVING.

3. Preparing the assignment

With your lab partners, write a full report with subsections of Introduction, Methods, Analysis, and Conclusions. See the rubric on Canvas for items that you need to address in your report. You can also use the sample report (under Files/Lab handouts/Sample Lab Report (Empirical Modeling).pdf on Canvas) as a guide.

Before leaving the lab today, discuss the following:

- When should you meet to prepare the report? Recall that we strongly discourage you from dividing up the tasks without discussing the final product.
- Who should do what tasks before the meeting?

BEFORE YOU GO:

- ***Ask TA to sign off.***
- ***Be sure to turn off the Oculus headset by holding the power button (on the right side of the headset) before putting them away in the box.***

Dear Colleague Letter: Research on the *Manifold*

I. PRELUDE

The Manifold lab takes place over three weeks, followed by one week of final presentations. It is framed as a game, but the storyline aligns with how physicists actually uncover new knowledge as a professional community. Your entire class will work in tandem to characterize a variety of novel phenomena in the VR lab, collaboratively forming a body of data and explanations for each phenomenon by generating and testing hypotheses.

II. INTRODUCTION

The *Manifold* is a recent discovery which has granted us entrance into myriad pocket universes, each with its own new forms of matter which obey unique rules for interacting with each other. The four fundamental force laws that govern our universe (see Fig. 1) are known to us, and the force laws of the matter in the *Manifold* seem to exist alongside them. Its discovery has prompted a scientific gold rush; though no matter can be brought back from the *Manifold*, it represents a font of knowledge no research scientist can resist. You are researchers tasked with understanding and cataloguing the force laws that govern the *Pockets* of the *Manifold*.

The National Science Foundation (NSF) supports the exploration of the *Manifold* and offers grants to promising research groups for cutting-edge measurement tools and expanded *Manifold* access.

III. SCHEDULE

Lab 4A: Conduct a hypothesis-generating study of the *Pocket* assigned to your group. Write up your findings as your *Pocket Exploration Grant Report*.

Lab 4B: Peer review each others' reports and make revisions. Write a *Pocket Analysis Grant Proposal* describing an experiment testing another group's hypothesis. This proposal will include a budget used to purchase new and upgraded measurement tools.

November 22: (Monday of Thanksgiving week) Turn in your revised *Pocket Exploration Grant Report* and your completed *Pocket Analysis Grant Proposal*.

Lab 4C: Respond to feedback on your proposal. Carry out your proposed experiment. Write up your findings as an oral presentation in the final week's *Symposium*.

IV. GRADING

The Manifold lab emulates the real-world experience of working as a researcher by focusing on developing knowledge of the *Manifold* through iterative experimentation by your entire class. Thus, effective communication of your findings to other groups is key to your success in this lab.

This lab, excluding the presentation, is worth 10% of your course grade. Here's how that will break down, out of 3 parts.

2 parts: *Pocket Exploration Grant Report.*

1 part: *Pocket Analysis Grant Proposal.*

Your final presentation during the *Symposium* will be worth 15% of your final course grade.

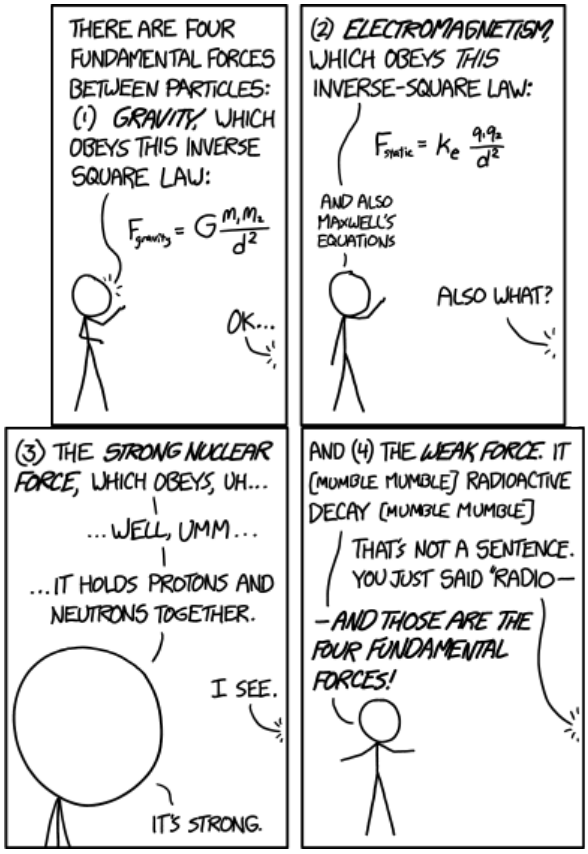


FIG. 1. A summary of the fundamental forces of our universe. Credit: XKCD

Manifold Lab Part A: Pocket Exploration

I. CONDUCTING YOUR INVESTIGATION

As first-time investigators of the *Manifold*, **your research group has been awarded a *Pocket Exploration Grant*** by the NSF. This grant comes with access to one *Pocket* of the Manifold, and the set of scientific equipment you are already familiar with: distance measurement tools, force meters, universal spatial anchors, and temporal perception adjustment modules (time-freezing powers).

As part of the *Pocket Exploration Grant*, **your group proposed to explore your pocket of the *Manifold*,** conduct an initial study of the matter found there, and report your results back to the NSF and your peer research groups.

This hypothesis-generating study of your *Pocket* is similar in nature to the Minty Particle lab. In your research group's proposal to the NSF, you promised to construct a model of the interactions between particles of the new matter you encounter, generating hypotheses that can be tested by other researchers.

II. REPORTING YOUR RESULTS

The NSF has specific requirements for what should be included when reporting your results. This report should be formatted according to the Overleaf template you have previously used, and satisfy the criteria described in the hypothesis-generating report rubric.

Here are some suggested report elements to include to get you started:

- Description of the distinct type(s) of matter found in your *Pocket*;
- Identification of physical relationships observed in the interactions between entities in your *Pocket*;
- Where appropriate, data describing these physical relationships, and representations of these relationships in mathematical form(s) suggested by the data.

As every *Pocket* is unique in its own way, the expected content and details of formatting of this report beyond the criteria above are yours to define. Each research group is the expert in the matter found in their *Pocket,* and can describe their *Pocket's* matter **in the way that they feel best reduces their data to succinct findings.** Reports may include videos, screenshots, etc. if appropriate.

In addition to the hypothesis-generating report rubric, there are some specific criteria your report must satisfy:

- Your report must offer a qualitative description of the behavior of your phenomena (how would you describe your observations in the first 30 seconds in the *Pocket*?)
- Another research group must be able to reproduce the experiments you document in your report.
- Your findings must be testable by your peer researchers (because you will be testing each others' findings).

III. NEXT STEPS

Upon submission of your *Pocket Exploration Grant Report,* the NSF sends it for peer review and requires authors to participate in the peer review of other reports. Based on the state of the report following peer review, the NSF program manager will decide whether to accept or return your submission.

To ensure adequate time for peer review, all *Pocket Exploration Grant Reports* are due midnight the day BEFORE lab.

Should the NSF return your *Pocket Exploration Grant Report* submission, they will provide instructions to help ensure its acceptance after revision.

Manifold Lab Part B: Pocket Analysis Grant Proposal

I. INTRODUCTION

Today, your research team will make further revisions to your *Pocket Exploration Grant Reports*. You will then be assigned one from another group, design an experiment to test their model, and write a grant proposal to acquire the funding to carry out that experiment with new and improved equipment.

Section II contains instructions for writing the proposal. Section III lists the additional equipment you may request in your proposal's budget.

II. POCKET ANALYSIS GRANT PROPOSAL

A. Select a Model

To support equal access to *Manifold Pockets* and reproducibility of results, your proposed *Pocket Analysis Grant* project must focus on another *Pocket* previously investigated by another research group's *Pocket Exploration Grant* project, building upon the hypotheses described in their report.

There will be printed copies of each group's report in the lab room for you to browse. You may factor multiple groups' findings on the same *Pocket* into your proposed experiment if you so wish. Take your time coming to a consensus on which *Pocket* your group wants to analyze. If your class has more than three groups, it is a good idea to take the time to skim the reports you did not peer review to inform your choice. Your TA will take your group's preferences into account when assigning your *Pocket* to analyze.

B. Proposal Requirements

Once you have selected a *Pocket Exploration Report* to build upon, write your proposal. Your proposal must include the following sections:

1. Introduction

The hypothesis must be described in detail; what model of the *Pocket*'s matter did the first group develop? You may need to do some interpretation and distillation of the findings in their report. This section doubles as your "background" section, where you describe the existing body of knowledge underlying your work. Include a detailed and succinct description of the model of the *Pocket* proposed by the *Pocket Exploration Report*(s) you are referencing.

Relevant rubric items:

B1: Identify the phenomenon to be investigated
B1a: Tone: Talking to a scientifically literate audience that is unfamiliar with this experimental set-up
B1b: Framing: What is the scientific question motivating this investigation?
C1: Is able to identify the hypothesis to be tested.

2. Methods

You should aim to identify multiple "directions" for investigation of the *Pocket*, if possible, nominally by digging into different elements of the proposed model or by considering multiple potential experimental methods. Briefly describe each of those, select one (or multiple, only if feasible), and justify your choice. Expand upon your selection in this section:

- A description of the possibility(ies) for further exploration of the *Pocket,* building upon the model. In other words, what are the different experiments you could plausibly conduct to test the model?
- A choice of which avenue to explore in your *Pocket Analysis Grant* project:
  - What hypothesis is your group testing?
  - What experimental procedure will you use to test this hypothesis?
  - What new equipment (see Section III) will you request from *REMI* to facilitate this experiment? Justify your choice. How will your new equipment support your experimental design?
  - Assuming the hypothesis holds, what is the predicted outcome of your experiment? Be specific. What analysis method(s) will you use to obtain this outcome?
  - If your experiment produces results inconsistent with the hypothesis, what will that look like?

Evaluation of this section will incorporate satisfactory answers to the above questions as well as your proposal's adherence to a few rubric items you are familiar with:

B3: Decide what physical quantities are to be measured and identify independent and dependent variables
C2: Is able to design a reliable experiment that tests the hypothesis
C4: Is able to make a reasonable prediction based on a hypothesis
C4a: Assumptions explained
C4b: Prediction follows from hypothesis
C4c: Prediction based on procedure

3. Budget

In this section, list the equipment you are requesting from *REMI*. In the Methods section, make sure to describe in detail how that equipment will enable your group to perform your experiment. You may request up to four (4) credits' worth of new and improved equipment.

Your proposal MUST be submitted by midnight on Monday, November 21, 2022.

III. REMI CATALOGUE

Reality Exploration Measurement Instrumentation (*REMI*) is a close partner of the NSF, and home to the world's leading experts on manifesting scientific equipment from our universe in the *Manifold*. The NSF's *Pocket Exploration Grants* come with a *REMI* equipment requisition, allowing the grantee to request development of new equipment for use in the *Manifold*.

REMI has a number of projects in the works, which can be completed and made available to your group through acceptance of a *Pocket Exploration Grant Proposal.* If you require equipment or functionality not described below, contact your NSF program manager to discuss a custom equipment requisition.

Each accepted *Pocket Exploration Grant Proposal* will be allocated 4 *REMI credits* to spend on upgrades prior to carrying out their experiment.

Force meter:

1 credit: Improve precision by 1 decimal place.

2 credits: Ability to lock rotation.

2 credits: Ability to "latch" to highest recorded value.

Measuring tape:

1 credit: Improve precision by 1 decimal place.

2 credits: Measuring tapes will "snap" to alignment with the principal X, Y, and Z axes of the workspace.

Manipulator:

1 credit: Add a velocity readout that gives the velocity with which the object you are holding is being moved.

2 credits: *(requires previous upgrade)* Unlock the option to specify the velocity with which an object is moved toward or away from you with the thumbstick.

Pause Physics:

2 credits: Pause physics after a predetermined amount of time.

Workspace Dimensions:

1 credit: Quintuple the width or depth of your *Pocket*'s workspace.

IV. NEXT STEPS

Your proposal will be reviewed prior to your next lab period. If your program manager identifies significant flaws in your plan, you will be invited to address them at the beginning of the next lab prior to acquiring your new tools and proceeding with your experiment.

Manifold Lab Part C: Pocket Analysis Grant Project

I. INTRODUCTION

Today, your team will carry out the experiment you proposed in your *Pocket Analysis Grant Proposal.*

II. POCKET ANALYSIS EXPERIMENT

A. Follow Your Procedure

You wrote your own instructions for this section; follow them! You may be asked to revise small parts of your proposal at the beginning of lab; note these in your lab notebook.

Take detailed notes on your experiment in your lab notebook as you proceed through your experiments. Remember that it is normal in science to revise and reconsider an experimental procedure as new information comes to light.

III. NEXT STEPS

During lab next week: Your group will give a 10 ± 2 minute presentation on your *Pocket Analysis Grant* findings, aimed at someone with a solid technical/scientific background but no knowledge of the work you have been doing in this course. Find the rubric for this presentation on Canvas.

By the end of lab next week: You will submit your presentation slides for grading. The slides must be content-complete and self-contained; any information you give verbally that is not on the slides must be present in the speaker's notes.

Appendix C

PERC PROCEEDINGS

Leveraging virtual reality for student development of force models in the introductory lab

Jared P. Canright, Jack R. Olsen, and Suzanne White Brahmia
Department of Physics, University of Washington, 3910 15th Ave NE, Seattle, WA, 98195

Providing students with opportunities to develop models in the introductory physics lab is made difficult by the dearth of physical phenomena that are both accessible to students at the introductory level and aren't easily found with a Google search. In response to this difficulty, we have developed a set of laboratory activities that leverage the unique affordances of virtual reality to provide a learning opportunity in which introductory physics students engage in the authentic practice of physics model development for a novel phenomenon. We present preliminary evidence of expert-like modeling behaviors, both in the context of hypothesis generation and in the context of modeling a measurement tool, as well as suggestions how virtual reality labs can be engineered to target other laboratory learning outcomes. Further, we discuss difficulties associated with this approach to curriculum design and ways they might be mitigated.

2020 PERC Proceedings edited by Wolf, Bennett, and Frank; Peer-reviewed, doi.org/10.1119/perc.2020.pr.Canright

I. INTRODUCTION

Modern virtual reality (VR) technology offers unique affordances, leveraging stereoscopic 3D rendering and embodied input modalities to create immersive learning environments that offer new ways to enrich student learning in physics [1]. Despite the possibilities VR enables, or perhaps in part due to their sheer variety, quantitative evidence of VR improving physics learning outcomes remains elusive. Prior research has compared the effectiveness of a VR intervention with the effectiveness of equivalent desktop computer-based and/or hands-on interventions at teaching introductory physics content. Application of this methodology to Moon phases [2], electrostatics [3], and magnetostatics [4] has found that students generally prefer VR treatments, but their enthusiasm is not accompanied by measurable learning gains over other treatments.

We describe here a pilot study of a different paradigm of VR use in the physics classroom. Rather than applying VR as another medium through which to teach existing physics content, we seek to employ it as a means to engineer physical phenomena and the tools available to examine those phenomena. In doing so, we can target specific physics laboratory learning outcomes in an instructional context where those outcomes are otherwise difficult to achieve.

We specifically seek to enable the authentic practice of physics model generation in an introductory electromagnetism laboratory. This goal is motivated by difficulties moving away from "verification" labs in high-enrollment introductory courses, which have been characterized as "exercises that verify known answers to unexciting questions" [5] and criticized for failure to measurably improve students' understanding of physics content [6] or foster expert-like views of experimental physics [7]. Physics model generation is recognized as a key learning goal of undergraduate physics laboratories by the AAPT [8] and is the cornerstone of the Investigative Science Learning Environment (ISLE) [9], a research-validated instructional system that develops expert-like beliefs and habits of mind about experimental physics.

The dearth of physical phenomena that are accessible at the introductory level for which students cannot readily find the "right" answer presents challenges for development of inquiry labs. Introductory students often consider any known model to be a standard by which their data is measured, rather than a conjecture subject to refinement by experimentation. Stein, Smith, and Holmes found that even in an inquiry-focused lab curriculum, at least 30% of student groups engaged in questionable research practices in attempts to make their data (incorrectly) confirm the model under investigation [10]. Hu, Zwickl, Wilcox, and Lewandowski conducted a qualitative study into students' belief that physics experiments can be completed without an understanding of the involved physical concepts and mathematical relationships. The authors concluded that, "...incorporating some lab activities for which the outcome is not known...might have a significant impact on students' understanding of the importance of experimental physics as a mechanism for uncovering new physics and driving the creation of new theoretical models." [11]

The Novel Observations in Mixed Reality (NOMR) lab project addresses this challenge by presenting representations of familiar and fictitious physical phenomena in an interactive 3D virtual learning environment (VLE) experienced via VR. Introduction of novel physical phenomena known only to the developer completely eliminates students' ability to identify an existing model or other "source of truth." Their sole source of information about novel phenomena is their own observations. A model-generating NOMR lab activity was pilot-tested with students enrolled in an introductory electromagnetism course. This paper seeks to document:

- The design of the NOMR labs, the phenomena explored within them, and the tools available to explore those phenomena;
- Preliminary evidence of expert-like modeling behavior employed in students' work in a NOMR lab;
- Discussion of opportunities to employ this approach to target other laboratory learning outcomes;
- Difficulties associated with this approach and strategies for their mitigation.

II. CONTEXT AND METHODS

A. Instructional Context

This Winter 2020 study took place in the honors introductory electromagnetism course at the University of Washington (UW), the second term of a three-quarter course where ISLE-based labs are being developed as part of the UW Laboratory Transformation Project. The lab sections met for two hours weekly. The course enrollment of 26 collaborated as eight persistent groups, four groups in each of two lab sections.

The 1-week NOMR lab activity that is the focus of this paper was the final lab of the quarter. Earlier in the term, students participated in a 2-week NOMR lab. Students quantitatively tested a model for interactions between VR analogues of electrically charged particles, and then developed and tested a model (Coulomb's law) for interactions between unfamiliar, fictitious "minty" particles. Minty particles carry a single charge type which repels itself at close range and attracts at long range, obeying a force law that resembles Coulomb's law with an additional constant negative term. Students developed strategies for using virtual measuring tools to gather force and separation data to develop a model, and became adept at using the VR technology. In the rest of the paper, we report on the later NOMR lab, in which students built on these skills.

B. Virtual Learning Environment

We refer to the software package running on the VR system in which students carry out the lab activity as the virtual learning environment (VLE). We developed the VLEs described below in the Unity game engine [12]. The VLE provided the particles for consideration, distance and force measurement tools (shown in Figure 1), and other capabilities:

Force meter: Force meters analogous to familiar spring scales consist of a pivot anchored at a point in space, a bob to which a single particle can be attached, and a rigid rod of fixed length connecting the pivot and bob. When a particle is attached to the bob, it swings around the pivot in response to forces upon the particle, coming to rest where the unit vector $\hat{r}$ pointing from the pivot to the bob is parallel to the force experienced by the particle $\overrightarrow{\mathbf{F}}$. The meter displays a scalar value equal to the dot product of these quantities: $F = \overrightarrow{\mathbf{F}} \cdot \hat{r}$, expressed in arbitrary force units. This can be thought of as the scalar force felt by a rigid spring connecting the particle and the pivot, with positive and negative values representing extension and compression of the spring, respectively.

Measuring tape: Two small spheres connected by a line accompany a display that gives the distance in meters between the spheres.

Anchors: Particles can be anchored, fixing their position at a point in space and preventing them from moving in response to collisions or field-mediated interactions.

Toggling particle interactions: A "Toggle Physics" button allows the user to freeze/unfreeze all particles in the scene, allowing for configuration of experiments.

Paint: Particles all start out green regardless of their identity, and can be colored yellow, purple, black, or white as a visual aid.

We refer to a collection of distinct *particles* in the VR lab and the relationships which govern their interactions as a *scenario*. All particles' charge type (if applicable) and magnitude are immutable, and all have the same immutable mass. Each particle produces a field $\overrightarrow{\mathbf{E}}_{type}$ associated with its charge type q_{type}. Each field independently obeys the superposition principle.

Each of the four groups within a lab section explored a different scenario; we report on two scenarios in sections III and IV. Students were informed that each scenario had no more than three distinct particles.

C. Laboratory Setup and Directions

Each group of two to four students was provided two identical setups, each consisting of a desktop computer and an Oculus Quest [13] VR system. While a student operated the VLE, the VR headset display was mirrored to the desktop computer via `scrcpy` [14] such that their partner(s) could see what was happening and record observations. Students were prompted to take turns using the headset.

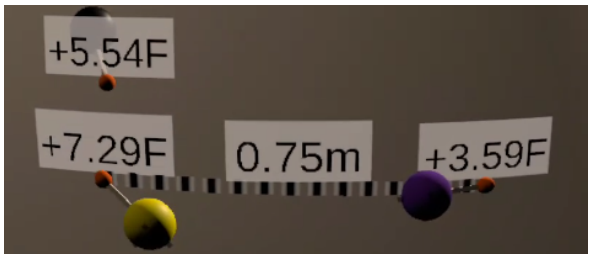


FIG. 1. This screenshot from the Triforce scenario shows power (yellow), wisdom (black), and courage (purple) particles interacting, each attached to a force meter. A measuring tape displays the distance between the pivots of the force meters attached to the power and courage particles.

Each student received the same lab procedure handout at the start of the lab; none were previously briefed on the lab content. Students were first directed to develop observational experiments to answer the questions:

- How many types of particles does the VR lab generate?
- How can the identity of a particle be determined?

Once students identified the number of particle types present in their scenario, named them, and developed a procedure for identifying them, they were directed to identify the independent variable(s) relevant to their particles' interactions and develop an evidence-backed mathematical model relating one such variable to the force between particles.

D. Research Methodology

While this paper is meant largely as a presentation of a unique curriculum design, case studies of two student groups' activities during the lab are presented in the following sections. Video data of select groups' lab activities were recorded and transcribed, and enhanced by field notes taken by the first author. These data were used to reconstruct a picture of notable moments in students' work during the lab activity.

III. TAGALONG SCENARIO

This scenario's only particle (the tagalong) produces a field dependent upon the relative velocity of the source particle $\overrightarrow{\mathbf{v}}_{src}$ and the velocity of the inertial reference frame $\overrightarrow{\mathbf{v}}_{frame}$ in which the field is measured, using Galilean relativity:

$$\overrightarrow{\mathbf{E}}_{tag}(\overrightarrow{\mathbf{v}}_{src}, \overrightarrow{\mathbf{v}}_{frame}) = q_{tag,src}(\overrightarrow{\mathbf{v}}_{src} - \overrightarrow{\mathbf{v}}_{frame}).$$

A tagalong with charge $q_{tag,i}$ experiencing a net tagalong field $\overrightarrow{\mathbf{E}}_{tag}$ as measured in its rest frame experiences a force proportional to the field:

$$\overrightarrow{\mathbf{F}}_{tag,i} = q_{tag,i}\overrightarrow{\mathbf{E}}_{tag}.$$

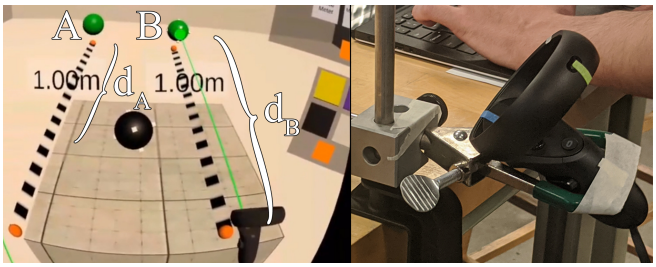


FIG. 2. Left: The Scouts' experimental setup in the VLE, with $N = 1$ black anchored tagalong between free tagalong A and manipulated tagalong B. Right: The Oculus Quest controller taped to the lab stand held B on a fixed path. With this configuration, moving B a distance $d_B = 1$ m toward the controller at a constant velocity causes A to move $d_A \approx 0.5$ m in the same direction before coming to rest.

This force causes a system of tagalongs to tend toward a state where its linear momentum is distributed equally among all tagalongs. Thus, if one tagalong in a moving system is anchored, the linear momentum of the system quickly goes to zero as every other tagalong slows to a halt.

A pair of students whom we will refer to as the Scouts worked through the Tagalong scenario. Through exploration of the behavior of a system of many tagalongs in response to manipulation of a single tagalong, the Scouts developed an initial model: "Particles accelerate to match the average velocity of particles around it." Casting this verbal representation into mathematical form results in the relationship implemented to govern this scenario: $\overrightarrow{\mathbf{F}}_{tag,i} \propto \overrightarrow{\mathbf{v}}_{avg}$. In the process of developing a quantitative experiment, the Scouts determined that every particle behaved the same way, and that their interactions were independent of relative position.

To test this model, they proposed to anchor N tagalongs, leave one tagalong (A) floating freely, and move another tagalong (B) a specific distance d_B at a constant velocity $\overrightarrow{\mathbf{v}}_B$. Thus, they reasoned, the velocity to which the free particle would try to accelerate is given by $v_{avg} = v_B/N$ for the time particle B is in motion $t = d_B/v_B$, after which v_B and v_{avg} return to zero. The Scouts' model predicted that the distance traveled by the free particle d_A would be proportional to v_{avg}, and thus inversely proportional to N: $d_A \propto 1/N$. The Scouts set out to measure d_A for various N to test this hypothesis.

When setting up this experiment, the Scouts recognized that they needed a means by which to keep d_B and $\overrightarrow{\mathbf{v}}_B$ constant between trials. The VR controllers can move a particle away from or toward the controller in a straight line by moving the controller thumbstick. The Scouts suspected (and confirmed with a TA) that pushing the thumbstick all the way up or all the way down caused the held particle to move at a constant velocity v_B. The Scouts found a very creative way to move B along the same path every trial, keeping consistent d_B and $\hat{\mathbf{v}}_B$: They taped a controller to a lab stand as in Figure 2 to fix its position for the entire experiment, never letting go of B.

The experimental design and configuration of apparatus performed by the Scouts demonstrates NOMR labs' potential to create an opportunity for students to exercise incredible creativity in experiment design in an introductory lab. However, their accurate model was backed by a flawed analysis. Their assumption that $v_{avg} = v_B/N$ should have been $v_{avg} = v_B/(N+1)$ to account for the influence of particle B. While the Scouts recorded d_A for $N = 0$ and $N = 1$, these data were absent in the Scouts' plot of $1/N$ versus d_A, which they presented to support their ultimate conclusion that $d_A \propto 1/N$.

IV. TRIFORCE SCENARIO

The Triforce scenario features three unique particles, each with its own charge type and associated field; these are dubbed power, wisdom, and courage. For brevity, we describe their interactions as a set of force laws governing interactions between pairs of particles, all of which satisfy Newton's third law. The force exerted on a particle of type 2 by a particle of type 1 is given by

$$\overrightarrow{\mathbf{F}}_{1\leftrightarrow 2} = k_{1,2} C \frac{\hat{r}}{r^2},$$

where C is a constant of proportionality independent of particle type and $k_{1,2} = +1,\ -1$, or 0 for attractive, repulsive, or zero force, respectively, as given for each combination of particle types in Table I.

TABLE I. Interactions between pairs of Triforce particles.

1 \ 2	**Power**	**Wisdom**	**Courage**
Power	Attraction	Repulsion	Attraction
Wisdom	Repulsion	Attraction	Zero
Courage	Attraction	Zero	Repulsion

A group of three students to whom we will refer as the Trio worked through the Triforce scenario. The process of determining the number and properties of the particles present in this scenario was more involved due to the presence of three unique particles. This process was confounded by a software bug that teleported a courage particle behind a wall in the VLE, where it remained, attracting other power and courage particles. The rogue courage particle was only discovered and deleted after a TA was asked why particles were being mysteriously attracted to a wall:

TA: "I mean, maybe the force is asymmetric, but..."
Student 1: "...we could reset and try this again."
Student 2, in headset: "I don't like the idea of a force being asymmetric. That, like, breaks the second law."
Student 3: "Wait, the force shouldn't be asymmetric, should it?"
Student 2: "Well, it shouldn't, but this is a simulation, so..."

Student 1: "Should we try resetting, just to make sure we don't have any stray particles?"
Student 2: Agrees and resets the VLE.

This exchange shows that students briefly entertained the notion that the particles' strange tendency to be attracted in a certain direction was a result of deliberate violation of Newton's laws in the VLE. While this is a promising example of students' willingness to consider a variety of explanations for a phenomenon, it also demonstrates that VLE bugs are dangerous. Had the Trio not reset the VLE, this bug may have led them down a frustrating rabbit hole.

With the bug resolved, the Trio soon arrived at the conclusion that their scenario had three unique particles and successfully identified the nature of the interaction between each pair of particles, constructing Table I. At this point, we observed that by giving a three-student group a scenario with three unique particles, we laid the groundwork for an unanticipated expert-like behavior: The students named the particles after themselves.

In their presentation, the Trio presented quantitative evidence for a $\overrightarrow{\mathbf{F}}_{p\leftrightarrow p} \propto \frac{\hat{r}}{r^2}$ attractive force law between two power particles, and suggested that other pairs obey the same law but noted that this claim was untested.

V. DISCUSSION

This work demonstrated a pilot physics laboratory curriculum design which leverages the unique affordances of VR to enable authentic physics model development at the introductory level. The groups who worked with the Tagalong and Triforce scenarios both demonstrated expert-like modeling behavior, identified using the ISLE scientific ability rubrics for observational experimentation [15]. Both clearly identified and described the salient features and patterns of their scenario in words and connected that representation to a data-backed mathematical relationship that accurately characterized parts of their scenario.

In light of successful observational experimentation, the Scouts' trip-up in their testing experiment raises interesting questions. They found that measurements of d_A for $N = 2 \ldots 7$ fit a believably linear trend in a plot of d_A versus $1/N$, whereas $N = 0, 1$ did not. To resolve this discrepancy, the Scouts omitted $N = 0, 1$ in that plot. It is unclear why the Scouts found fault in those data, rather than their analysis or their model. Video data suggest that the Scouts (rightly) believed their model $\overrightarrow{\mathbf{F}}_{tag} \propto \overrightarrow{\mathbf{v}}_{avg}$ was well-supported by qualitative observational experiments. This suggests that the Scouts were unable or did not attempt to resolve the discrepancy by examining the relationship between their model and their derived experimental prediction $d_A \propto 1/N$. Further study of students' use of questionable scientific practices in absence of a "correct" model may aid development of inquiry-based lab curriculum.

This trial suggested that NOMR labs may enable other laboratory learning goals:

Modeling measurement tools: We were concerned that students may misinterpret the force meter's signed scalar readout of a vector quantity, but it turned out to be a robust tool with just enough depth to require careful thought but not obstruct students' ability to conduct experiments. Examination of student modeling of virtual measurement tools through a framework which treats that process as a partner to modeling a physical system [16] may inform the design of NOMR labs which intentionally build this skill.

Designing experiments: Given the same set of measurement tools, the Tagalong scenario required a much more complicated experimental design than Triforce. By careful consideration of what tools are available in a NOMR lab and how they must be combined to make sense of a phenomenon, one could regulate the complexity of experiments expected of students.

We also find that a VR lab is not without its difficulties:

Implementation and facilitation of VR in the classroom comes with significant monetary and technical overhead, and requires special TA training and protocols to ensure safety of students using the headsets and those around them.

VLE creation currently requires significant VR development experience, which is generally not expertise readily available in a physics department. By exploring this approach further, the necessary elements of physics VLE authoring tools may be identified, paving the way for development of authoring tools designed to be used by physics educators.

VLE physics bugs are an epistemological hazard. As found in Section IV, unintended physics bugs can be interpreted as intentional behavior by students. If their VLE had included the ability to create uniform global fields or charge walls, the bug may have gone unnoticed for the duration of the lab. Comparing how students think about troubleshooting VR tools and physical tools is a promising area for further study.

These findings beg the question: What unique advantage does VR lend to the NOMR labs over desktop simulations? We have a sense that VR produced effects that could not be replicated on 2D screens, but this study does not aim to make that argument. Our experience is consistent with Greenwald's findings that direct connection of sensorimotor inputs with dynamic simulations, blending of multiple representations, and reduced cognitive load are all likely contributors to any unique benefit associated with VR physics curriculum [1]; future work will draw from his methodology to provide a more concrete answer to this question.

ACKNOWLEDGMENTS

The authors thank Robert Pecoraro and Charlotte Zimmerman for their invaluable effort, and gratefully acknowledge financial support from the National Science Foundation GRF (DGE-1256082), ARCS Foundation, AAPT Bauder Fund, and UW Student Technology Fee.

[1] S. W. Greenwald, *The Equipped Explorer: Virtual Reality as a Medium for Learning*, Ph.D. thesis, MIT (2018).

[2] J. H. Madden, A. S. Won, J. P. Schuldt, B. Kim, S. Pandita, Y. Sun, T. J. Stone, and N. G. Holmes, Virtual Reality as a Teaching Tool for Moon Phases and Beyond, in *PERC Proceedings* (Washington, DC, 2018).

[3] J. R. Smith, A. Byrum, T. M. McCormick, N. Young, C. Orban, and C. D. Porter, A controlled study of stereoscopic virtual reality in freshman electrostatics, in *PERC Proceedings* (Cincinnati, OH, 2017) pp. 376–379.

[4] C. D. Porter, J. R. Brown, J. R. Smith, E. M. Stagar, A. Simmons, M. Nieberding, A. Ayers, and C. Orban, A controlled study of virtual reality in first-year magnetostatics, in *Physics Education Research Conference 2019*, PER Conference (Provo, UT, 2019).

[5] R. Trumper, The Physics Laboratory - A Historical Overview and Future Perspectives, Science and Education **12**, 645 (2003).

[6] N. G. Holmes, J. Olsen, J. L. Thomas, and C. E. Wieman, Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content, Physical Review Physics Education Research **13**, 10.1103/PhysRevPhysEducRes.13.010129 (2017).

[7] B. R. Wilcox and H. J. Lewandowski, Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics, Physical Review Physics Education Research **13**, 10.1103/PhysRevPhysEducRes.13.010108 (2017).

[8] J. Kozminski, H. Lewandowski, N. Beverly, S. Lindaas, D. Deardorff, A. Reagan, R. Dietz, R. Tagg, M. Eblen-Â-Zayas, J. Williams, R. Hobbs, and B. Zwickl, *AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum*, Tech. Rep. (AAPT, 2014).

[9] E. Etkina and A. V. Heuvelen, Investigative Science Learning Environment - A Science Process Approach to Learning Physics, in *Research-Based Reform of University Physics*, Vol. 1 (2007).

[10] M. M. Stein, E. M. Smith, and N. G. Holmes, Confirming what we know: Understanding questionable research practices in intro physics labs, in *PERC Proceedings* (Washington, DC, 2018).

[11] D. Hu, B. M. Zwickl, B. R. Wilcox, and H. J. Lewandowski, Qualitative investigation of students' views about experimental physics, Physical Review Physics Education Research **13**, 10.1103/PhysRevPhysEducRes.13.020134 (2017).

[12] Unity Technologies, Unity Real-Time Development Platform.

[13] Oculus Quest.

[14] scrcpy Github page.

[15] E. Etkina, A. Van Heuvelen, S. White-Brahmia, D. T. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and A. Warren, Scientific abilities and their assessment, Physical Review Special Topics - Physics Education Research **2**, 10.1103/PhysRevSTPER.2.020103 (2006).

[16] B. M. Zwickl, D. Hu, N. Finkelstein, and H. J. Lewandowski, Model-based reasoning in the physics laboratory: Framework and initial results, Physical Review Special Topics - Physics Education Research **11**, 10.1103/PhysRevSTPER.11.020113 (2015).

Developing expertlike epistemologies about physics empirical discovery using virtual reality

Jared P. Canright and Suzanne White Brahmia
Department of Physics, University of Washington, 3910 15th Ave NE, Seattle, WA, 98195

The process and role of developing new scientific models experimentally is a cornerstone of physics – and arguably an important learning objective of laboratory courses – but research in physics education suggests that the content-focused, "cookbook" paradigm of lab instruction common in the last 2-3 decades does not effectively meet this objective. These instructional practices often leave students believing that in-class experiments exist to supplement conceptual learning and that the sole purpose of real-world experiments is to test theories. As a result, many students frame lab work as an exercise in knowledge confirmation, and often develop the expectation that all physics that appears in the labs can be found with a Google search. Framing a real-world lab such that its results are Google-proof enough to engage students in the process of scientific discovery requires great creativity. This paper describes an alternative approach, using virtual reality to integrate the process of authentic scientific discovery into the lab curriculum. We developed labs in which students explore and develop mathematical models for unknown force laws between new kinds of particles that we invented. We find that this intervention produced significant shifts in students' epistemologies about experimental physics toward being more expertlike – specifically regarding the central role experimentation plays for research physicists in making new discoveries and developing theories, and the role of laboratory experiments in teaching a nuanced and unique scientific way of knowing.

2021 PERC Proceedings edited by Bennett, Frank, and Vieyra; Peer-reviewed, doi.org/10.1119/perc.2021.pr.Canright

I. INTRODUCTION

Understanding how to shift students' epistemologies to reflect the importance of experimentation in the discovery of new physics is an open problem in physics education research on labs. Creation of new scientific models is an essential element of the scientific process [1], but students rarely get experience doing so in the "cookbook" style of introductory physics labs common in the last 2-3 decades [2, 3], many of which remain in use today. Improving the match between instruction and authentic science practices requires overcoming barriers:

- Great creativity is required to design lab activities in such a way that students don't see their findings as something that has been discovered before and can be Googled.
- Students overwhelmingly believe that the role of experiments in physics courses is to confirm something they already know [3, 4].
- While there are exceptions, physics lab curricula in large-enrollment courses generally do not reflect the iterative nature of scientific inquiry, leaving students with a poor understanding of the complex interplay between experiment and theory involved in the creation of new scientific models [3, 5].

In this paper, we document effects of a novel intervention on students' epistemologies about experimental physics. This intervention provides the opportunity to generate models of new force laws through experimentation using a virtual reality (VR) environment [6]. We code free-response survey data to answer the following research questions:

(RQ1) If students engage in experimentation as a process for theory development, how will it effect their epistemologies about experimental physics?

(RQ2) How do students compare their experience in labs that mimic experimental research to other kinds of laboratory experiences?

II. BACKGROUND

A. Student Epistemologies

To better understand student epistemologies about physics experimentation, we replicate a study reported on by Hu and Zwickl [7], hereafter referred to as the "original study", or "original." These researchers developed an instrument to explore beliefs about experimentation, theory, and their interplay held by physics lab students. They administered the instrument to physics students with a range of experience: 28 introductory calculus-based students in a combined studio/lab course, 20 upper-division undergraduates in standalone modern physics lab courses, and 31 PhD students. This instrument is composed of four open-ended questions, accompanied by a codebook used to identify certain themes in question responses in a reproducible way.

Q1 In your opinion, why are experiments a common part of physics classes? Provide examples or any evidence to support your answer.
Q2 In your opinion, why do scientists do experiments for their research? Provide examples or any evidence to support your answer.
Q3 In your opinion, what defines a scientific theory?
Q4 How do theory and experiment relate? Provide examples or any evidence to support your answer.

Introductory students in the original study exhibit an almost singular focus on the idea that experiments exist to supplement conceptual learning. This contrasts with upper-division undergraduates and PhD students, who acknowledge the role of in-class experimentation in the development of scientific abilities, and as a means to better understand the scientific process. Likewise, introductory students describe the relationship between experimentation and theory almost universally as experimentation serving to test or prove theories. More experienced students described the same relationship as an iterative, holistic process in which experiments can give rise to new theories, theories can explain experimental findings, and theories can guide the development of experiments.

We describe our use of Hu and Zwickl's instrument to assess the impact of a novel curriculum on students' perception of experimentation in physics in Section III C.

B. Hypothesis-Generating Curriculum

Our framework for designing curriculum to teach model generation is the Investigative Science Learning Environment (ISLE) framework [8]. In ISLE, model generation happens in *hypothesis-generating experiments*, where students engage in open-minded exploration with the goal of developing a model for an unknown phenomenon. This phase of the ISLE cycle, itself a simplified representation of the scientific process, is followed by testing the models created and further revising, refining, and applying them.

Hypothesis-generating experiments involving novel scenarios are difficult to create, especially in introductory courses. In many cases, experimental physics questions that could reasonably be investigated at the introductory level are well-known with easily-Googled answers. In presence of known answers, students tend to hold those in highest regard, seeking to confirm known answers above anything and everything else, even in the face of contradictory data [9]. Thus, any access to a "right answer" derails efforts to engage students in authentic model generation.

C. NOMR Labs

Our approach to this problem uses VR labs to simulate physical laws that do not exist in nature or the Internet. We include the constraint that they be consistent with known physics that actually does exist so that students can

rely on their physics knowledge when reasoning in the VR space. All of the "fictitious" physical laws in this intervention can be construed as hypothetical mathematical variations on Coulomb's Law.

We call these NOMR labs [6], short for Novel Observations in Mixed Reality. In NOMR labs, students explore fictitious physical phenomena in an immersive 3D environment. The virtual apparatus is designed such that it does not give perfect answers; experimental uncertainty is still very much present even in the simulation. The "right answers" programmed into the simulation are never shared with students, such that the only "right" answer is the one that students can make the best case for.

III. METHODOLOGY

A. Virtual Learning Environment

We refer to the VR application running on the headsets as the virtual learning environment (VLE). The VLE offers students a common set of tools in every lab, including measuring tapes, force meters, the ability to anchor individual particles or pause physics entirely, and a source of the particles they are experimenting upon. The NOMR VLE design is described in detail in a prior publication [6]; the VLE used in this study includes visual and quality-of-life upgrades over the previous version but no notable changes to its controls or capabilities.

B. Instructional Context

Our data come from the Winter 2021 administration of the University of Washington's honors introductory electromagnetism lab; 37 students, mostly freshmen, worked in persistent groups of 3-4, meeting weekly in 2-hour remote sessions. Five of nine weeks of lab activities were NOMR labs; the other four weeks went to a 1-week pen-and-paper uncertainty lab, a 2-week PhET-powered circuits lab, and a week of final presentations. NOMR labs were run remotely by students with Oculus Quest headsets (either their own or or borrowed from the Physics Department). Each group had one such headset operator, who streamed the display to the rest of their group over Zoom.

The first VR lab, called Charge and Mint, lasted two weeks. Groups designed and conducted an experiment to test whether virtual analogues of electric charges follow some re-scaled manifestation of Coulomb's law. Then, they took qualitative and quantitative data to create an empirical model for the interaction between fictitious *minty* particles, which behave according to an unknown force law.

The second VR lab, called the Manifold Lab, lasted three weeks. Students explored several sets of new types of particles that were discovered in the *Manifold*, each set obeying a unique fictitious physical force law. Each group was assigned at random one of four sets of particles. In the first week, they performed a hypothesis-generating experiment to create a testable model of the unique force law. Each group submitted a report describing their experiment and model to a class-wide repository.

During class in the second week, each group selected a report describing a model of a set of particles they had not yet seen. They wrote and submitted a proposal before the third week of lab, describing an experiment to test the other group's model. These experiments were carried out in the third week, and their results presented in a symposium in the fourth week.

The Manifold Lab was presented to students with a game-like narrative in which they function as research scientists. They explore different "pocket" universes with novel forms of matter obeying fundamental force laws unknown to our universe. Before performing the second experiment, students wrote a "grant proposal" that summarized another group's findings, proposed an experiment to test their model, and requested additional or improved equipment within the VR lab. Each group had a few points to spend on equipment, e.g. more precise measurement tools, a larger workspace, tools that "snap" to more convenient configurations, or the ability to automatically pause physics after a set amount of time.

This design seeks to emulate the experience of working within a scientific collaboration: The class as a whole collaborates by sharing data and designing experiments to and test and revise each other's models. The privilege to conduct the second experiment with better equipment depends on applying (non-competitively) for grant funds.

C. Research Methodology

We administered a survey following the course's final exam for a small amount of participation credit. This survey included the four questions from the original study [7] as well as a fifth question, presented in another section of the survey, asking students about their experience of the Manifold Lab:

Q5 The framing around Lab 4 (the "Manifold Lab") was designed to put you and your group in a setting that emulates the experience of working within a scientific collaboration. Did you feel that this improved or detracted from your experience? Why? What, if anything, would you change about it?

Responses to **Q1**-**Q4** were coded independently by the first author and by a colleague using the published codebook. After discussion and some context-specific clarification of our codebook we reached 99% inter-rater agreement across all codes. We present here on a subset of codes identified in responses to **Q1**, **Q2**, and **Q4**. In the interest of constraining the scope of our discussion to student views of experimentation, we do not present on the responses to **Q3** in this paper.

Responses to **Q5** are used to gauge students' perceptions of the Manifold Lab's similarity to their previous lab experiences and to real-world science.

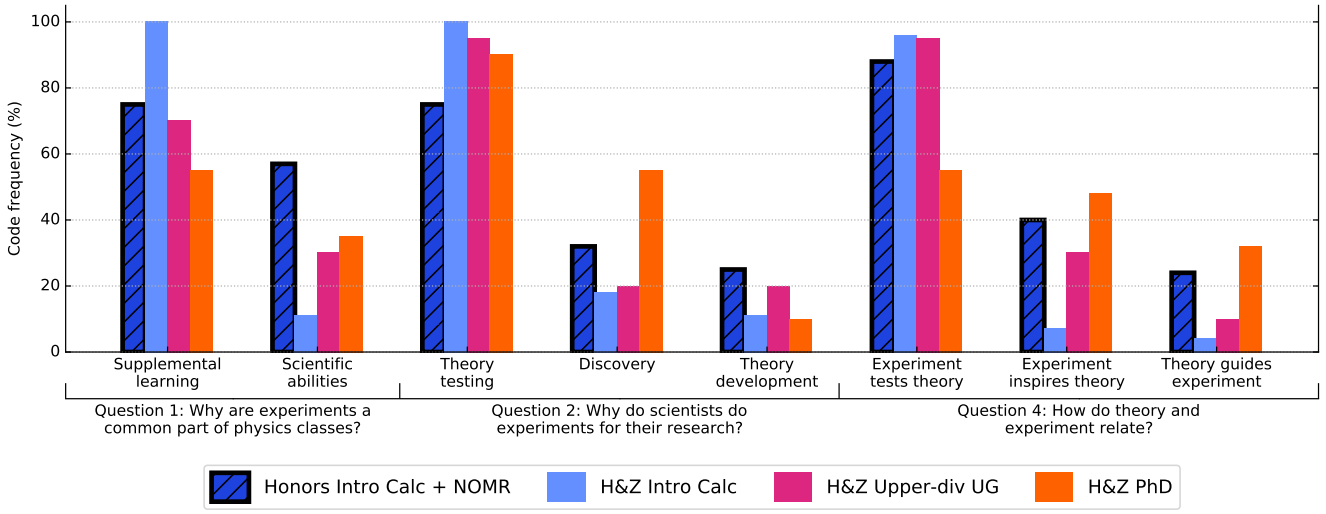


FIG. 1. These charts show the percentage of responses from each population of students that received each code. "H&Z" in the legend indicates data from Hu and Zwickl [7]. Cool colors represent introductory calculus students (the "novice" population) and warm colors upper-division undergraduates and PhD students (the "expert" population).

IV. ATTITUDES QUESTIONNAIRE

Results from our students are compared to reported data from the original study's introductory calculus-based students (the "novice" population) in Figure 1. We also show results from the original study's upper-division undergraduate and PhD student populations, which we consider to be the "expert" populations. We note that our honors intro physics students likely represent a more advanced population in terms of their scientific epistemology, perhaps falling in a space between the original's introductory students and their upper-division undergraduates.

We make quantitative comparisons of qualitative data, which is inherently less precise than quantitative data. We therefore refrain from making claims about their statistical significance until these effects can explored in greater depth [10]. We find the patterns to be interesting on their own.

Responses to **Q1**, *Why are experiments a common part of physics classes?*, were assigned up to two codes:

Supplemental learning: Experiments provide supplemental learning experiences for concepts and theories.

Scientific abilities: Experiments help cultivate students' scientific abilities, such as experimental design, data collection, and data analysis skills.

We find that the students in our study think about in-class experimentation with less of a singular focus on supplemental learning, and a stronger belief that labs are meant to develop their scientific abilities. Even compared to experts, they are more likely to talk about in-class experiments developing scientific abilities.

Responses to **Q2**, *Why do scientists do experiments in professional research?*, were assigned up to three codes:

Theory testing: The purpose of doing physics experiments is to prove a theory or test a hypothesis.

Discovery: Experiments help investigate unknowns.

Theory development: Experiments inspire the development or improvement of theories.

Compared to intro students in the original study, we found that the students in our study have less of a singular focus on theory testing than any other population. Further, they mention discovery and theory development at a rate comparable to experts in the original study, well above that of intro calculus students.

Responses to **Q4**, *How do theory and experiment relate?*, were assigned up to three codes:

Experiment tests theory: Experiments are conducted to test, support, or back up a theory.

Experiment inspires theory: Experiments can inspire the development or modification of theories.

Theory guides experiment: Theories often provide the basis for the design of experiments.

We again find that students in our study do not universally believe that experiments exist just to test theories. We see more frequent description of experiment-inspired theory development and in theory as a basis for experimental designs. Comparing to the original study's experts, students in our study discussed the latter two ideas at a rate solidly between the two expert categories, upper-division undergraduates and PhD students. With the assumption that PhD students' responses are informed by experience working in research labs while the upper-division undergraduates' are generally not, these data suggest that the NOMR labs had a similar (but weaker) effect upon the epistemologies of students in our study as real-world research experience.

V. MANIFOLD LAB FEEDBACK

Student responses to **Q5** were positive nearly across the board. Complaints spoke to the organizational and administrative hiccups commonly associated with pilot testing a new set of activities, rather than to the content and design of the Manifold Lab. The most negative response described the experience as unremarkable:

> "It was fine. It felt somewhat like a normal lab."

Many responses remarked on the Manifold Lab's similarity to the real-world practice of experimentation:

> "...having our group test hypotheses, collect data by doing tests, and creating our own model for novel physical phenomena was a **great way to teach the process of "doing science,"** especially by emphasizing the importance of connecting the data to math. I also enjoyed the process of needing to test further with better equipment (and needing to save up for said equipment), which is also a **realistic situation** that happens in the real world."

> "...It really felt **like what I imagine it feels like to work in the academic community,** and was unlike any lab I had done previously. In particular, the concept of working with completely new phenomena without any previous model was a really fantastic experience..."

Students described significant changes to their thinking about lab activities, brought about by the absence of a known answer:

> "...we had to do a lot of troubleshooting and ended up with wild results, which I think was beneficial, because it required us to **reflect on our process and the uncertainty of our experiment** more."

> " Being able to work with a pocket universe and using experimentation to describe it was the best experience in physics courses I've ever experienced. I preferred this VR experience to any physical lab for the sole reason of it being entirely new and **having to get every ounce of info about it through experimentation and collaboration.**"

> "...we were able to **explore physics in the pockets instead of focusing on getting a grade for an assignment;** in addition, we could **design our own experiments** instead of following a set list of instructions which made it enjoyable. In essence, it felt more like an experiment than other labs which was nice."

Overall, students in our study responded positively to the game-like narrative structure of the Manifold Lab and shared the belief that it emulates the real-world practice of experimental physics to a greater degree than lab curriculum they had previously experienced.

VI. CONCLUSIONS

These findings raise the question: Was VR strictly necessary to achieve this outcome? We predict that embodied cognition and VR's improved similarity to hands-on experimentation as opposed to "flat" desktop simulations play a key role. Without a greater understanding of differences in students' perception and experience of VR and desktop simulations, we are not yet in a position to answer this question.

A comparison of the original study's results with ours revealed notable features of the epistemologies of students who experienced NOMR labs. The view held by nearly all of the introductory students in the Hu and Zwickl study - that the purpose of experimentation is to provide supplemental conceptual learning - is less commonly held by the students in our study. Many did not mention supplemental learning at all in response to **Q1**, instead focusing on the development of scientific abilities.

The students who completed NOMR labs exhibit a more nuanced understanding of the role of experiments and of the iterative relationship between experiment and theory that underpin real-world science. They acknowledge the role of experimentation in the investigation of unknowns and the development of new theories (**Q2**) and identify more facets of the interplay between experiment and theory (**Q4**) with frequencies comparable to experts.

This gives us an answer to RQ1: Engaging students in experimentation as a process for *theory development* may shift their epistemologies about experimental physics to reflect a more expertlike understanding of the complex role of experimentation in development and improvement of theories.

Feedback from the Manifold Lab suggests that students perceived it as an authentic simulation of the process of scientific theory development. Some reported that this approach gave them an unprecedented feeling of connection with real-world science, and had a transformative effect on how they think about experimentation as budding scientists. These insights answer RQ2: Students generally believe they learned more about the practice of science in activities that mimic experimental research than in their previous laboratory experiences.

We have presented findings suggesting that immersion in a novel universe as a research scientist on the cutting edge can shift students' epistemologies about scientific theory development toward expertlike views, reflecting their role in the narrative. With or without VR, we hope this approach aids instructional designers in tackling theory development as a learning goal in their lab curriculum.

ACKNOWLEDGMENTS

The authors thank Charlotte Zimmerman and Jack Olsen for invaluable support. This work was funded by the National Science Foundation GRF (DGE-1256082) and the UW Student Technology Fee.

[1] J. Kozminski, H. Lewandowski, N. Beverly, S. Lindaas, D. Deardorff, A. Reagan, R. Dietz, R. Tagg, M. Eblen-Â-Zayas, J. Williams, R. Hobbs, and B. Zwickl, *AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum*, Tech. Rep. (AAPT, 2014).

[2] R. Trumper, The Physics Laboratory - A Historical Overview and Future Perspectives, Science and Education **12**, 645 (2003).

[3] D. Hu, B. M. Zwickl, B. R. Wilcox, and H. J. Lewandowski, Qualitative investigation of students' views about experimental physics, Physical Review Physics Education Research **13**, 10.1103/PhysRevPhysEducRes.13.020134 (2017).

[4] P. H. Sneddon, K. A. Slaughter, and N. Reid, Perceptions, views and opinions of university students about physics learning during practical work at school, European Journal of Physics **30**, 1119 (2009).

[5] E. M. Smith and N. G. Holmes, Best practice for instructional labs, Nature Physics , 1 (2021).

[6] J. P. Canright, J. R. Olsen, and S. White Brahmia, Leveraging virtual reality for student development of force models in the introductory lab [in review], in *PERC Proceedings* (2020).

[7] D. Hu and B. M. Zwickl, Examining students' personal epistemology: the role of physics experiments and relation with theory, in *Physics Education Research Conference* (American Association of Physics Teachers (AAPT), 2018) pp. 11–14.

[8] D. T. Brookes, E. Ektina, and G. Planinsic, Implementing an epistemologically authentic approach to student-centered inquiry learning, Physical Review Physics Education Research **16**, 10.1103/physrevphyseducres.16.020148 (2020).

[9] M. M. Stein, E. M. Smith, and N. G. Holmes, Confirming what we know: Understanding questionable research practices in intro physics labs, in *PERC Proceedings* (Washington, DC, 2018).

[10] D. Hammer and L. K. Berland, Confusing Claims for Data: A Critique of Common Practices for Presenting Qualitative Research on Learning, Journal of the Learning Sciences **23**, 37 (2014).