Summer 2023

Teaching Non-Experimental Chemistry Content Through the Pairing of Strategically Curated Data Sets and Model-Based Inquiry

Andrew Banker

Teaching Non-Experimental Chemistry Content through the Pairing of Strategically Curated Data Sets and Model-Based Inquiry

by

Andrew Banker

A capstone project submitted in partial fulfillment of the requirements for the degree of

Masters of Arts in Teaching

Hamline University

Saint Paul, Minnesota

May 2023

Capstone Advisor: Julianne Scullen, Ed.D.

Content Reviewer: Sarah Hick, PhD

Peer Reviewer: James Carlson

DEDICATION

To my best teachers in my early years as a teacher: the 2021 Harding summer school students, the 2021-2022 Harding general chemistry, accelerated chemistry, and IB chemistry students, and the 2023 Roseville summer academy students. I made this because of the spark I saw in your eyes.

ACKNOWLEDGEMENTS

Thank you Backstory Coffee Roasters in Saint Paul for housing my best hours of academic thinking and writing. Special thanks to my most influential academic strategist, brainstormer, challenger, and hype-builder, Emily. Thanks for the nudge to get started on this. Special thanks to my north star, Aurelia. You keep me grounded and soft. Thank you to Sam DiVita for your collaboration and encouragement in my early years of teaching. Finally, immense gratitude to my mentor, Sarah Hick. I would not be built for sustainably being a teacher without your voice in my life.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER ONE

Introduction

Aiming to Synthesize Content and Process

I could not care less about every citizen in the USA learning about chemistry content. If I were king for a day, I would replace the state graduation requirement with a required course about evidence standards or health epidemiology in a heartbeat. My Minnesota 9 - 12 Chemistry license is something I treasure, but I regard it carefully. The value of getting to teach chemistry lies, for me, in the opportunity that the discipline of chemistry offers to practice the scientific way of constructing knowledge and grappling with uncertainty. As a set of theories about matter, energy, and change, high school chemistry is merely a foundational class for science majors in college. Yet as a training gym for analyzing data, assessing evidence and probing uncertainties, a chemistry course is an empowering opportunity to own the strengths of scientific thinking for whatever arena of influence you pursue.

As I have experienced my educational philosophy shift towards valuing scientific practices high above any disciplinary concept, I have changed my perspective about what I consider a meaningful use of class time. It matters more and more to me that students feel empowered to utilize a way of knowing than for them to know anything in particular. In science education there has been a long conversation about whether to emphasize teaching the "process" of scientific knowledge construction, or the "content" of particular science disciplines (Millar & Driver, 1987). One could say that I have swung further to the "process" side of this supposed content-process dichotomy (Millar & Driver, 1987). Nonetheless, in the momentum of science education in America, there is no foreseeable disappearance of chemistry as a common science course taken by high-schoolers. Furthermore, there are reasonable arguments for a more balanced synthesis of the content-process dichotomy that I find at least realistic and useful in the context of 21st century education in America (Passmore et al., 2009; Millar & Driver, 1987), even if I am not epistemologically convinced of the merits of teaching chemistry as the particular content of choice. What I am motivated to establish is this: *What kind of curricular resources effectively synthesize the parallel goals of engaging students in scientific practice and understanding disciplinary content?* To orient this question in the American high school context, I provide some background on the current shift to the Next Generation Science Standards (NGSS), my personal shift to a student-centered pedagogical paradigm, my experimentation in my own approach to classroom teaching, and my deliberation of a research trailhead that can answer my primary question.

The Backdrop of the NGSS in Minnesota

The Next Generation Science Standards (NGSS) are the basis of the 2019 Minnesota science standards currently being adopted and implemented across Minnesota (MDE, n.d.a). This set of standards make a marked shift from the timbre of the 2009 Minnesota State Standards implemented by the Minnesota Department of Education (MDE). The 2009 Standards had an overwhelming emphasis on the chemistry concepts that were to be taught (MDE, 2010). The 2009 Standards did list a number of *Nature of Science* standards that were meant to be explored throughout the various science courses in high school, but did not assign any particular course the task of integrating them (MDE, 2010). These *Nature of Science* standards hosted much of the language about the science practices that I became so passionate about. However, it is easy to imagine how many schools and teachers could, not maliciously, end up neglecting these standards in pursuit of the disciplinary ones. In contrast, the NGSS systematically integrates science and engineering practices (SEPs) standards with a more limited set of disciplinary core ideas (DCIs) standards historically prioritized in science courses (NGSS Lead States, 2013). By making this serious attempt to elevate the importance of SEPs, I hope the NGSS and the 2019 Minnesota science standards can contribute to a tide-change in how science is taught.

Living Into a Student-Centered Pedagogy Paradigm

Even though my educational training gave me many tools to pursue a more student-centered paradigm in my pedagogy, translating this paradigm to the real context of schools proved to be quite difficult. I felt the heavy pressure of teaching all the content standards, creating reasonable cohesion with others teaching the same course, and the cultural impulse to teach how I was taught as a student. My shift to being able to live into a student-centered pedagogy paradigm began when I was able to consider its potency for making science learning relevant to my students.

Early on in my educational training, I became keenly aware of the importance of the term *relevance*. Teachers were supposed to aspire to do whatever they could to make their content as *relevant* as possible to their students. As a science teacher, I wrestled with this immensely. It was difficult to conceptualize how to make electron energy levels or chemical naming conventions *relevant* to students' lives. My attempts to bridge this relevance gap in my content consistently felt inauthentic and forced.

During my science methods courses, though, I was able to broaden my definition of what counted as *relevant*. The term itself harkens to the concept of relating. I began to understand that *how* we approached learning in the classroom could have just as much potential to help students relate to the learning as *what* we were actually learning. The most ground-breaking revelation for me was realizing that if you started learning a concept by eliciting students' questions, observations, or critiques, the concept could all-of-a-sudden be transformed into something *relevant*, without the inauthenticity of superficial connections to real life. Shifting the nature of the exploration to be more student-centered in this way could build in relevance to my science teaching, even if the content itself was initially quite alien to the students.

Eventually, I began to see the ways that I could weave multiple threads into a stronger cord. Student-centered lessons lent themselves to true scientific inquiry. Student-centered learning could amplify the relevance of the science content; engaging students in the practices of science could make the learning inherently more student-centered; the practices of science could be an empowering set of skills that are relevant beyond the science content itself. Combined, I imagined these interrelated priorities could significantly strengthen each other.

Exploring New Approaches in the Classroom

Once I was teaching in the classroom, I had my opportunity to implement such a dynamic, but it took a while to get there. Due to the nature of student teaching and the debut of the Covid-19 pandemic, it wasn't until summer school ahead of my third year of teaching that I was actually able to experiment with my own curricular approach *in* a physical classroom. The building where we conducted the summer school credit recovery program was at an urban high school with no air conditioning. Fans blared loudly to circulate air. In the summer school setting that year, there was a chemistry curriculum vacuum that I had to fill in order to teach my students. I began to try using strategically curated data sets (think *examples arranged to show patterns*) as the foundation of every day's chemistry lesson. We got out half a dozen 2 ft. x 3 ft. whiteboards. *What do you notice here?* became the guiding question every day. I progressively developed a sense of probing questions that kept the ball in their court as they would make sense of the patterns in the data sets. Slowly, rituals and routines began to emerge that helped the students wrestle with the data sets and build competency around the underlying concepts. I had thought that summer school was destined to be boring, but I came home genuinely energized about it each day. We co-created a vibrant learning community in that sweaty, no-AC, urban school building. I knew I had struck some sort of gold mine, but there was much more to explore.

The following school year I set the intention to explore how these strategically designed data sets could be the foundation of engaging in scientific practices like hypothesis generating, model development, data analysis, model refining, collaboration, communicating conclusions, and addressing uncertainties. It became clearer and clearer that what we were attempting together in my class was to model the scientific process akin to how it is pursued at the highest levels. We were a thread of that larger tapestry of science. It felt so fitting, so purposeful, so real. In the larger world of science, there is no teacher to tell you "Correct!" when you propose a hypothesis or a model. You must test your idea on examples to see if it is replicable. You must test it further to see if it is fully generalizable to all contexts. You must identify and consider possible rationales for the rules and the exceptions to the pattern. You must compare your articulation of your models to those of other scientists and come to a consensus about what the most convincing and evidenced models are.

The process of wrestling with data sets in this way was a radical shift from the all-too-common paradigm of *science is knowable facts; the teacher will tell you what they are*. Not only were students learning about electron configurations, but they were doing so by practicing the skills of noticing patterns, creating models, discussing them, and refining their thinking as they applied their ideas to novel examples. I began to wonder if we were traversing territory that the Next Generation Science Standards mapped in their envisioning. I saw us engaging in the science and engineering practices (SEPs) in the process of discovering the disciplinary core ideas (DCIs) of chemistry (NGSS Lead States, 2013).

Defining the Research Scope

In my chemistry classes that year, we were building the airplane as we flew it. That meant there were imperfections in our model, as one would expect from any burgeoning model. I worked closely with my Peer Assistance and Review mentor, Sam DaVita, to refine aspects of this burgeoning approach. It was becoming apparent by this point that, while I had stumbled onto a promising and rewarding trailhead, that I had many questions to consider:

1. What portion of the success of this process was due to the design of the process?
2. What portion was due to the repeated practice in our rituals and routines?
3. What portion was due to the attitudes and growth-mindset messaging in the class?
4. What portion was subjectively due to my personality and presence?

5. What portion was due to the students themselves, many of which were advanced chemistry students?
6. Is there any variation in results based on the level of cohesion subjectively experienced by students in different groupings or classes?
7. Why were some students blossoming with this approach, while others remained disengaged?
8. Why were my colleagues struggling to create the same classroom culture when they tried implementing my data sets and approach in their classrooms?

These questions in their entirety may not be able to be completely answered until research on similar approaches to science teaching is more thoroughly conducted. Nonetheless, I decided that I wanted to review the literature underlying the concepts I was exploring and create a chemistry curricular resource that synthesized the evidence base into a collection of data sets and an accompanying protocol that chemistry teachers would find useful and justified. While I set out to create an example unit in my content area of chemistry, my hope was that my example project could serve as a proof-of-concept that science teachers of various disciplines could pattern their curricular approach after. Many of the high school science courses seek to teach the same scientific and engineering practices for their disciplinary content. I also knew I would have to be open to changing my thinking as I plumbed the depths of the related literature. At onset I could not assume that all aspects that I had experimented with were positively contributing to the overall positive results I was observing. I narrowed my many questions into a more focused research question that I could answer as a trailhead in my journey towards greater clarity and understanding. The question I pinpointed for my project was: *What kind of curricular resource effectively synthesizes the parallel goals of engaging students in scientific practice and understanding disciplinary content?* I believe answering this question will be a timely contribution to the field of science education, particularly to current and preservice teachers in Minnesota as they grapple over the next few years with the shift to the NGSS' emphasis on infusing disciplinary content teaching with engagement in scientific and engineering practices. Without example blueprints of how this shift can be pursued, I worry that many teachers will spend their energy trying to simply justify how their current approaches already accomplish the goals of the NGSS well enough. A compelling, attainable, useful, and evidence-based curricular resource could be a foot in the door for many educator's shift to a more student-centered pedagogical paradigm in science education.

In the next chapter, I present the pre-existing conversation in the academic literature on inquiry based approaches to science learning, the empirical evidence base for inquiry based learning, a more detailed description of the NGSS, and the particularly useful approach of model-based inquiry (MBI). This sets the stage for justifying my capstone project: a MBI-inspired protocol paired with strategically designed data sets reflecting non-experimental chemistry content.

CHAPTER TWO

Literature Review

Introduction

While many secondary science educators theoretically value scientific processes and inquiry in their curricular approach, there can be difficulty imagining what those processes look like when teaching students the disciplinary content also laid out in the standards. *What kind of curricular resource effectively synthesizes the parallel goals of engaging students in scientific practice and understanding disciplinary content?* Teachers feel the tug between their ideals for science teaching and the need to *just keep going* to make it through all the content standards. There seems to be an unfortunate belief that scientific practices and disciplinary content teaching are at odds, and that the teacher can only prioritize one or the other. The spiral of anxiety may sound like this: *if we focus on engaging in scientific practices we will use up too much classroom time; we will never get to the topics of acids and bases within the school year; these students will be at a disadvantage in their college chemistry course; it will be all my fault*. The intention behind this capstone project was to create a useful chemistry curricular resource that convincingly pairs scientific practices with disciplinary content learning, rendering such concerns less potent for science teachers.

This capstone project presents a model-based inquiry (MBI)-inspired protocol paired with strategically designed data sets reflecting non-experimental chemistry content. While this curricular resource is focused on chemistry content, the MBI-inspired protocol is meant to be more widely useful in all content areas of science education at the secondary level, needing only discipline specific data sets to be paired with it. Data sets in this project are considered to be any graphical, pictorial, mathematical, linguistic, physical, simulational, or model-based set of information that can be *mined* for patterns and relationships. The MBI-inspired protocol presents a structure for how the teacher may sequence classroom conversations and activities in a way that supportively guides students in their inductive thinking regarding the data sets. Together, these two curricular resources give specific form to the goal of engaging students in scientific practice and understanding disciplinary content. To provide the necessary context for this project, this review of the literature discusses the history and evidence base behind inquiry-based instruction overall and specifically details the scientific practice of modeling as exemplified in MBI. Exploration of these topics establish the theoretical and empirical base for many of the choices made in the design of the proposed curricular resource. To begin with, the role of inquiry in science education is probed.

Inquiry in Science Education

Inquiry has been an important topic in the discourse of science education for more than sixty years. Its prominence can be traced back to the materials developed by the Biological Sciences Curriculum Study (BSCS) in the early 1960's (Anderson, 2008b). From there, *inquiry* became a widely used term in science research, including being extensively referenced in the science standards and rationale developed at the turn of the century by the National Research Council (NRC) (1996; 2000). The role of inquiry in science education has been especially detailed in the publications of Roger Bybee, who chaired the content working group for the *National Science Education Standards* (1996) and later became executive director for the BSCS (Bybee, 2002).

Definition of Inquiry

In its usage in the *National Science Education Standards* (NSES), the term *inquiry* was meant to draw direct parallels to the scientific processes utilized by researchers in the various fields of science (Anderson, 2008b; Bybee, 2002; NRC, 1996). Inquiry was meant to evoke the components of the scientific journey of asking questions, developing appropriate methodologies, collecting quality data, analyzing that data, and communicating conclusions in terms of the evidence for them. While this characterization is a reasonable one, there indeed are many different conceptualizations of what inquiry means (Anderson, 2008b). For example, some emphasize how the nature of inquiry traces back to fundamental human curiosity and a trial-and-error approach to testing predictions and explanations (NRC, 2000). As an imprecise word, it is just as important to establish what one means by the term as it is to use it. Common threads that synthesize many impressions of the term *inquiry* as used in this literature review are as follows:

1. It an extension of natural human curiosity
2. It is a student-centered, active learning process
3. It involves the implementation of different processes of mining information from data sources in order to discern meaning
4. It reflects various practices of scientists

There are other aspects not included above that are featured in some definitions of inquiry. For example, many consider inquiry to be necessarily hands-on and related to laboratory experiments (Bybee, 2002; Colburn, 2000; Lonergan et al., 2019). For the sake of being able to apply greater generalizability in the use of the term, other possible components of a definition of inquiry are not discussed here. Instead, the exploration of the literature on inquiry here focuses on various models of inquiry and the empirical data probing its usefulness.

Levels of Inquiry

Four Levels of Inquiry. A number of writers have modeled inquiry in terms of *levels*. Most notably, Banchi & Bell (2008) provided descriptions for four levels of inquiry: confirmation inquiry, structured inquiry, guided inquiry, and open inquiry. These forms of inquiry can be conceptualized as being on a gradient scale with a high degree of teacher design on one end (confirmation inquiry) and a high degree of student design (open inquiry) on the other end. The four levels of inquiry are summarized below, with confirmation inquiry denoted as Level 0 on the inquiry scale as Abrams et al. did to clarify its non-inquiry nature (2007):

> Level 0: Confirmation inquiry is where the answer to the question of interest is already known by the students, but they carry out a predetermined experiment or activity designed to confirm the explanation already provided to them. This approach is labeled as Level 0 to indicate that it does not actually incorporate the student involvement that is foundational to what *inquiry* is. Hereafter this approach will be referred to as a *verification lab* so as to not imply that it is a form of inquiry.
>
> Level 1: Structured inquiry is where the question, methods, and materials are all provided by the teacher, but the results are unknown prior to the students engaging in the experiment or activity.
>
> Level 2: Guided inquiry is where the teacher presents the research question and provides materials but does not prescribe the methods to be implemented in

pursuit of answering the research question. Students must design their own methodology.

Level 3: Open inquiry is where the students are responsible for defining their research questions, developing a methodology to answer the question, and carrying it out themselves. This highest form of inquiry is closest to that of formal science research. (Banchi & Bell, 2008)

Learning Cycle. In contrast to Banchi & Bell (2008), Colburn (2000) did not include confirmation inquiry in his list of inquiry-based instruction in his article *An Inquiry Primer*. Colburn similarly described structured inquiry, guided inquiry, and open inquiry but included an additional form of inquiry called a *learning cycle* (2000). In their iteration of the concept, a learning cycle involves three major phases: exploration, content, application (Colburn & Clough, 1997). They detailed that the exploration phase resembles a guided inquiry, with the teacher providing a question of interest and the materials needed to explore a phenomena or concept. Then, as they explain, in the content phase, the teacher explicitly names and explains the concepts that the students discovered in their initial inquiry. They purport that the verbal explanation is more relevant and less abstract to students after having the opportunity to initially explore a context to which the concept is bound. Finally, they discuss how the application phase is where students are asked to use their newfound understandings to solve a novel situation.

Three-Dimensional Model for Guidance in Inquiry. Vorholzer and von Aufschnaiter (2019) provided a striking three-axis model to synthesize choices that can be made in designing inquiry experiences for students. One axis of their model describes the degree of autonomy afforded to students: low to high. The second axis describes the degree of conceptual information: minimal, implicit, or explicit. The third axis denotes the cognitive domain of guidance: content or procedural (Vorholzer & von Aufschnaiter, 2019). This model is particularly valuable for considering possible scaffolding routes when sequencing student experiences of inquiry. It can also be an accountability tool for teachers to assess where the activities they design fall in this three-dimensional model. To illuminate a possible example, a guided inquiry process may give students moderate autonomy (first axis), minimal *or* implicit conceptual information (second axis), and procedural guidance, but not content guidance (third axis).

Inquiry Learning as Constructivist Learning

Applied to the context of the classroom, inquiry seems to have much resonance with constructivist learning (Anderson, 2008a; Zuckerman et al., 1998). Constructivist learning found its roots in the work of Soviet psychologist, Lev Vygotsky. Vygotsky and Cole (1978) claimed that learning was inherently social, where the student's individual abilities can be expanded by them practicing concepts in their zone of proximal development (ZPD) with assistance and interaction. Through this process of knowledge and ability construction, the student gradually envelopes the previously unattainable schema into their individual competency (Vygotsky & Cole, 1978). Anderson (2008a) relied heavily on a constructivist framework in his characterization of inquiry learning, noting how inquiry learning and constructivist learning both entail meaning being actively constructed by students in a social context. Zuckerman et al. emphasized how inquiry as an individual skill could be rehearsed and practiced through sociocultural communication and cooperation (1998). The processes of developing questions, methods, models, and forms of scientific communication are key aspects of inquiry that lend themselves to social constructivist learning.

Constructivist science learning included but transcended the focus of the conceptual change paradigm, particularly in offering a framework that better addresses the achievement gap in science learning (Anderson, 2008a). The conceptual change framework assumes a more Piagetian understanding of learning, where a process of internal cognitive progressions is assumed as a student engages with the material world. The conceptual change tradition was concerned with how concept sequencing and addressing misconceptions could lead to more accurate understandings of concepts being developed by students. Constructivist science learning, on the other hand, was more Vygotskian in nature, even while its goals included students changing their thinking. There was a focus on how information was processed and practiced socially and linguistically among the people in the room. Science knowledge construction was assumed to be a process of social discourse and meaning-making. Because the social constructivist framework considered the learners in their cultural and sociological context, it gave credence to the real disadvantages students faced when there was a cultural mismatch between the language and interactions in school and their own cultural context. Constructivist teaching alerted the teacher to the opportunities to connect scientific discourse with the existing social language used by students as well as plan interactions that gave students adequate time to rehearse their burgeoning understandings through communication (Soysal, 2020). Nonetheless, there was also notable friction in the discussions around whether constructivism was a helpful term to describe the nature of science (Anderson, 2008a).

Even with certain arguable merits, constructivism received its own criticism as well. In 2006, Kirschner et al. set off a lively debate in the literature over the validity of inquiry-based and constructivist teaching with their provoking article titled *Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching*. The primary arguments presented discuss the inefficient use of cognitive load, inattentiveness to the differences between novice learners and experts, and the high cost of leaving misconceptions unchecked in the process of learning (Kirschner et al., 2006). Kirschner et al. suggested that novice learners get a chance to observe worked examples as a way to become familiar with the process and content being taught (2006). Later in 2006, Hmelo-Silver et al. directly responded to the article by Kirschner et al., claiming that the straw man characterization of inquiry-based and constructivist learning as *minimally guided* was reductionistic and incorrect (Hmelo-Silver et al., 2007). They argued that much inquiry-based research incorporates scaffolding precisely to mitigate the effects presented by Kirschner et al. They stress that students often did receive guidance in the inquiry process (except in fully open inquiry) and the teacher took on a very active role in pushing students to refine their thinking in the process. They also added that models of expert performance were often presented, and post-inquiry clarification lectures could address any residual misconceptions. Up to that point, the discussion of inquiry had been largely rhetorical and theoretical, although both sets of authors did point to some empirical evidence in their articles. Moving beyond that theoretical debate, it is prudent to now examine the empirical evidence base explicitly rather than solely rely on the summaries of writers and thinkers in this space.

Empirical Analysis of Inquiry as an Evidence-Based Practice

There have been a number of empirical studies probing the effect size of implementing inquiry-based science learning. Many writers are fervent that inquiry is securely an evidence-based practice (Blanchard et al., 2010; Bunterm et al., 2014; Geier et al., 2008; Hmelo-Silver et al., 2007; Rahman & Lewis, 2020). Some studies, though, suggest mixed results (Liou, 2021). Still others reject the paradigm of inquiry-based teaching as ineffective altogether (Kirschner et al., 2006). The following sections will discuss the empirical evidence base for inquiry used in the classroom.

Detroit Public Schools Study. One sizable study by Geier et al. involved approximately 5000 Detroit public schools (DPS) students in middle school that were either taught with inquiry-based instruction or standard DPS instruction (2008). By the metric of performance on the Michigan Educational Assessment Program (MEAP) science assessment, they listed an overall positive effect size of the inquiry-based instruction to be 0.44 and 0.37 standard deviations respectively in the two cohorts analyzed. They determined the effect size to be statistically significant. The size of this study and the breadth of the implementation across multiple schools gives this study arguable generalizability. Furthermore, the authors addressed many potential sources of sample bias, with convincing quantitative accounts for why the possible sample biases could not account for any considerable portion of the effect size (Geier et al., 2008).

An important limitation in the interpretation of these results was that the comparison groups were different in ways other than the style of teaching. The teachers that implemented the inquiry-based instruction received extensive professional development and received some ongoing support by graduate students and peer teachers

(Geier et al., 2008). This intentional support, because it was not applied in equal scale to the comparison group of general DPS teachers, may be a source of residual unmeasured confounding. While it seems unlikely to account for the entirety of the effect size, it brings up questions about what portion of the effect size was due to the nature of the inquiry-based instruction, and what portion of the effect size was due to the nature of having extensive professional development and support, regardless of the type of intervention.

Meta-Analysis of Evidence-Based Instructional Practices. Rahman and Lewis looked at 94 experimental or quasi-experimental studies in their metaanalysis of evidence-based instructional practices (EBIP) used in chemistry classes (2020). Their analysis included the strategy of process oriented guided inquiry learning (POGIL), making its analysis relevant to this review. The overall effect size they reported, measured as a corrected difference between the two groups analyzed in each component study, was 0.62. They considered this difference a medium to large effect size, with a confidence interval ranging from 0.522 0.713. This result justified their claim that any of the EBIPs may be used with some confidence in positive effect..

While the study aimed to be able to quantify the relative effect size of the different strategies analyzed, the authors admitted that conclusions about the relative effectiveness of each strategy could not be determined with much confidence due to confounding variables, particularly the variable of the test measure varying as a single-topic assessment or a cumulative assessment (Rahman & Lewis, 2020). Analyzed as a possible moderator of results, the studies with cumulative assessment demonstrated a much more moderate effect size (0.25) than the studies with single topic assessments

(0.87), indicating that this was a meaningful moderator of effect. In their analysis, the POGIL strategy had a weighted mean effect size of 0.30, but the 95% confidence interval touches the null [0.00, 0.60] (Rahman & Lewis, 2020). This was the only EBIP to span the null, but POGIL was also one of only two strategies that had a considerably larger median setting size, calculated as 109.5 (Rahman & Lewis, 2020). There was no definitive secondary analysis or correction for how setting size moderated the results, but the authors discussed the inverse relationship between setting size and effect size, indicating this variable as another meaningful moderator of results that would need to be controlled for more direct comparisons (Rahman & Lewis, 2020).

In regards to the generalizability of these results to inquiry in chemistry education, there are a few notes of concern. The meta-analysis included studies in secondary and postsecondary settings, meaning that the effect size could have also been moderated by this difference. In their short discussion of it, the average effect size seemed larger in the secondary setting (0.87) than in the postsecondary setting (0.50), but this was likely confounded by other features such as class size and postsecondary filtering effects (Rahman & Lewis, 2020). Another note of concern here is that inquiry was only represented expressly in one of the interventions, POGIL, while there were many other inquiry approaches that could have been included. In characterizing the effect of inquiry, then, the possible conclusions from this analysis must be conservative. Only 10 studies using POGIL were analyzed (Rahman & Lewis, 2020). The results for POGIL only were highly moderated by assessment coverage type, with a 0.87 effect size for single-topic, and 0.15 effect size for cumulative (Rahman & Lewis, 2020). These point estimates, however, were determined by a small number of studies, making meta-analysis difficult since variation in the individual studies could have ended up having a considerable effect on the final point estimate leading to very noisy data (Rahman & Lewis, 2020). Of no fault of the authors of the meta-analysis, there is substantial need for more experimental studies of the POGIL strategy to make significant and generalizable claims.

Inquiry Versus Teacher-Directed Instruction Taiwan. One sizable study from Taiwan raised important questions about what inquiry-based instruction did and did not accomplish in their schools. Liou's study (2021) analyzed data from the Program for International Student Assessment (PISA) results for a representative sample of 7,708 15-year-old students from Taiwan who either experienced inquiry-based instruction or teacher-directed instruction. The PISA served to give results about science achievement, but also to survey the students about contextual variables that they experienced in their classrooms (Liou, 2021). While the observational nature of this study lended itself to important limitations such as residual confounding and analytical flexibility, the sample size was a considerable strength that many other experimental studies did not possess and the choices in their methodology did have convincing strengths. Instead of assigning students to receive inquiry-based instruction or teacher-directed instruction, they established latent constructs for inquiry-based instructional practices, teacher-directed instructional practices, science self-efficacy, enjoyment of science, and instrumental motivation and then calculated factor loadings for each survey response with regard to its assigned latent construct (Liou, 2021). In the end, inquiry-based instructional practices had a clear negative effect on science achievement while teacher-directed instructional practices had a significant positive effect (Liou, 2021). While inquiry-based instruction had significant positive effects on enjoyment of science, instrumental motivation, and self-efficacy, teacher-directed instruction had similar positive effects on those measures as well, only somewhat diminished compared to inquiry-based instruction in the later two constructs (Liou, 2020). Liou's results should be taken seriously, but there are limitations in the design of this analysis.

The primary limitation was introduced in the choice to use student survey responses rather than a professional analysis to determine whether or not inquiry approaches were being implemented in the represented classrooms. This analytical choice may not have been substantiated as the author did not establish whether this is a reliable surrogate for inquiry-based instruction. The perception of the students was a questionable stand-in for professional determination of whether or not their teachers implemented inquiry-based instruction. It is also possible that the survey questions assigned to the construct of teacher-directed instructional practices may in fact have represented actions and experiences that were also present in classrooms where inquiry-based instruction was being implemented. These questions are unanswerable in this retrospective observational study. Despite these critiques, this study did illuminate a very real signal about how students' experience of teachers explaining scientific ideas, discussing student questions, and demonstrating ideas had significant correlation with their science achievement and attitudes towards science learning (Liou, 2020).

Guided Inquiry Versus Verification Lab. Blanchard et al. ran a large study designed to determine whether or not guided inquiry could supplant the status quo of verification laboratory instruction (2010). They designed it as a pre-, post-, and delayed post-test study design for 1700 students in classrooms of 24 different teachers based on a week-long laboratory-based forensics unit. They found that high school students in the guided inquiry group overall outscored those receiving traditional instruction, but there were confounding issues with the data, as it showed that the student groups were already fundamentally different before the intervention was applied. For example, their Figure 1 showed that the average pretest score in the high school traditional groups was already lower than that of the high school guided inquiry group (Blanchard et al., p. 597). In the overall scores, this critique has the potential to account for the effect size, while in other areas the time confounding was even more pronounced, making any conclusions very tentative. Likely, there are other covariates that make the teachers implementing guided inquiry different from the teachers implementing the traditional lab instruction. The authors acknowledged the role of the amount of professional development that the teachers received, for example, as a covariate which could explain why the confounding was observed. The authors could have chosen to analyze student scores based on the change in score as opposed to the absolute score. This sort of analysis could have corrected for the initial difference issues throughout the analysis. A true randomization protocol for choosing teachers for the intervention and running the study early in the school year would be two suggestions for addressing the confounding issues if a subsequent study is designed to replicate or build off of the results from Blanchard et al. (2010). An important insight that the study did bring forward is how the rate of free and reduced lunch (a surrogate for poverty level) among students at the school was a considerable moderator of test results irrespective of the instructional approach (Blanchard et al., 2010). Indeed their total posttest and delayed posttest scores were even markedly decreased from the pretest scores for high poverty high schools for both inquiry and traditional instruction. This can caution any science instructor to recognize the powerful effects of poverty on overall learning, and be a motivator to identify and implement both evidence-based science teaching practices, but also strategies that are paramount when working with high poverty student populations.

Structured Versus Guided Inquiry. As a final consideration of the empirical evidence regarding inquiry, it is worth discussing a study that examines the relative effectiveness of different levels of inquiry, using the model of inquiry levels. Bunterm et al. ran an experimental study of 239 students across 3 schools in Thailand to compare the relative effects of structured inquiry and guided inquiry, with guided inquiry being the form with greater student responsibility in the process (Bunterm et al., 2014; Banchi & Bell, 2008). The results indicated that guided inquiry was superior to structured inquiry for the measures of science content knowledge, science process skills, self-perceived stress, and scientific attitudes (Bunterm et al., 2014).

A strength of this study was how the curriculum was designed for use by the two comparison groups. They used the 5E Learning Cycle Model with the first three phases of Engagement, Exploration, and Explanation being where there were distinctions in the instruction between structured and guided inquiry groups, but then the final two phases of Elaboration and Evaluation phases were held constant for the groups (Bunterm et al., 2014). Since both groups were using a structure given by the research team, significant effects seen are more likely attributable to the type of inquiry as opposed to differences in the amount of support given to the teachers in the comparison groups. This is a strength that the Detroit Public School study was lacking (Geier et al., 2008).

While the results for the measures of science content knowledge and science process skills seemed convincing, there were issues with the interpretation of the effect on self-perceived stress and scientific attitudes. The results for self-perceived stress had a high variance based on which school was being analyzed, pointing to the likelihood that other unmeasured school-based covariates were moderating the results beyond the effect of the type of inquiry instruction (Bunterm et al., 2014). The results for scientific attitudes were also noticeably confounded because there are considerable differences in the scores of the two comparison groups already present in the pretest results (Bunterm et al., 2014). Furthermore, in the results for scientific attitudes, the Grade 7 groups had arguably superimposable results, making it difficult to assign any predictive or explanatory power to the type of inquiry instruction along this measure.

Summary of Empirical Evidence Base. It is apparent that there was some discordance in the available evidence about the effect size of inquiry-based instruction in the classroom. The studies examined are a non-exhaustive sample of evidence that could be discussed, so there is much more that could be argued for and against inquiry learning. After critical appraisal of these sources, the most evidenced claim to be made is that there needs to be more well-designed, large-scale, prospective, randomized, controlled experimental studies to answer questions about the effect size and the key moderators of inquiry-learning as an instructional approach. Nonetheless, there are key conclusions that can be made with some confidence. First, the effect of poverty in student achievement was of paramount consideration alongside evaluation of inquiry versus traditional instruction (Blanchard et al., 2010). Next, inquiry learning could be effective in improving student achievement even on standardized tests compared to the raw effect of existing educational practice, assuming that the teachers are receiving adequate support in implementation and continued consultation (Geier et al., 2008). Even deserving some critique, the study in Detroit Public Schools was the most authoritative evidence presented here to support implementation of inquiry-based instruction. Inquiry learning could also be more meaningful for science achievement and science process skills when students experience guided inquiry rather than structured inquiry (Bunterm et al., 2014). Process Oriented Guided Inquiry Learning (POGIL) could be implemented in chemistry classrooms with confidence that it was at least as effective as standard instruction, while humbly holding the likelihood that it was more effective (Rahman & Lewis, 2020). Lastly, the teacher implementing an inquiry process in the classroom must not neglect the importance of teacher explanation, answering student questions, and demonstrating ideas at some part of the process of learning in their classrooms (Liou, 2021). This caution reflects the warnings of Kirschner et al. about the lack of guidance in classroom learning, and was relatedly supported by the discussion of Hmelo-Silver et al. about the role of scaffolding and guidance in the inquiry process (Hmelo-Silver et al., 2007; Kirschner et al., 2006). A helpful model here might be the idea that it was equally important to determine what inquiry was *not* as it was to articulate what it *is*. Inquiry learning in the classroom need not be completely unguided or unsupported, and may very well need to be paired with a range of other strategies to be most meaningful (Hmelo-Silver et al., 2007). Further quality research on modulators of effective inquiry learning is called for to confirm what claims can be made more definitively.

Summary of Inquiry in Science Education

The importance of inquiry in the discourse around science education cannot be overstated. The inquiry reform paradigm has attempted to get students doing what scientists do in their various disciplines as they learn the science content in the K-12 standards (Anderson, 2008b; Bybee, 2002; NRC, 1996). Nonetheless, there has been much confusion around what constitutes inquiry and how to measure its effect in empirical research (Blanchard et al., 2010; Bunterm et al., 2014; Garcia-Carmona, 2020; Geier et al., 2008; Hmelo-Silver et al., 2007; Liou, 2021; Rahman & Lewis, 2020). This capstone project presents a curricular resource that is substantively inquiry-based. The chemistry content based data sets used with a MBI-inspired protocol explicitly involve students in the process of inquiry and the protocol takes into account nuances about teacher scaffolding and guidance as key features to ensure effective inquiry-based instruction. The next large area of interest to be reviewed is the scientific practice of modeling, as it is a useful focus when considering processes of inquiry that may or may not involve controlled experimentation (Passmore et al., 2009).

Modeling in Science Education

Model-based inquiry (MBI) has begun to enjoy some limelight in the discussion of science education in recent years, gaining its own carve-out in the literature on inquiry (Windschitl et al., 2008). Modeling, in a generic sense, is a practice where one examines phenomena, identifies features or parts of the phenomena, and represents how those features are related (Windschitl et al., 2008). Familiarity with modeling as a scientific practice is all the more relevant since the publication in the early 2010s of *A Framework for K-12 Science Education* (NRC, 2012), hereafter referred to as the *Framework,* and the *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013). Districts and schools across the US are in the process of adopting the NGSS. Since modeling was found as one of the eight scientific and engineering practices (SEPs) in the NRC's three dimensional model of science education, teachers can prioritize instructional strategies around modeling with the weight of the standards behind them (Dass et al., 2015; NGSS Lead States, 2013). In this review section, there will be an overview of the scientific practices detailed in the NGSS, a deeper discussion of modeling as articulated in the NGSS, and then an exploration of the larger range of sources that address modeling and model-based inquiry.

Scientific Practices

With the introduction of the NGSS, there has been a notable departure from using the term *inquiry* and a shift to using *scientific practices* instead. This reorientation is due to the ambiguous nature of what inquiry had come to mean pedagogically over its years of generous–and often incorrect–use (Garcia-Carmona, 2020; NRC, 2012). By defining exactly what types of practices are involved in scientific inquiry, there is less confusion in imagining and reimagining classroom activities to incorporate such practices.

The *Next Generation Science Standards* detailed eight scientific practices that were reflective of scientific inquiry:

1. Asking questions
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data
5. Using mathematics and computational thinking
6. Constructing explanations

7. Engaging in argument from evidence
8. Obtaining, evaluating, and communicating information (NGSS Lead States, 2013)

Omitted above is the language around engineering in the practices, as the focus of this review is focused on the scientific practices. The rationale for scientific practices set forth by the NGSS emphasized the value of having students do what scientists do and, thus, gain appreciation for the overlapping and various approaches to investigating, modeling, and explaining the world (NGSS Lead States, 2013). The *Framework* offered engagement in science practices as a way to avoid students getting the notion that science is just “a body of isolated facts,” but rather to equip students to become “critical consumers of scientific information” (NRC, 2012, p. 41). This theoretical framework resonated more with the sociocultural tradition of science education that assumes that science is a socially constructed practice of sense-making (Anderson, 2008a). It also paired well with the critical tradition in that it opens the door for students to understand how science can be at risk of ideological and institutional bias because of it being of human conception (Anderson, 2008a). The *Framework* also emphasized that scientific practices were useful for gaining deeper insight into the disciplinary concepts along with awareness of the social purposes of science (NRC, 2012). This was an important note since the focus on inquiry and science practices has been accused of becoming too dissociated from the scientific knowledge that it generates and perpetuating a cycle of singular focus on only practices *or* content (Millar & Driver, 1987).

Modeling in the NGSS

While each practice could be discussed at length, the primary scientific practice that undergirds the structure of this capstone project is the practice of modeling.

Nonetheless, the other practices absolutely overlap and coordinate with modeling in the process of scientific knowledge construction, so there will be brief discussion of other practices in the accounting of modeling here as well.

The *Framework* brought new emphasis to the role and importance of modeling as a core scientific practice (NRC, 2012). The *NGSS* listed diagrams, physical replicas, mathematical representations, analogies, and computer simulations as examples of models (NGSS Lead States, 2013). Any conceptual model was meant as a representation that is analogous to actual phenomena (NRC, 2012). The *Framework* characterized models as "a major means of acquiring new understandings" noting that they "identify key features" even as they are not a direct representation of reality (NRC, 2012, p. 78). Models are often iteratively improved as scientists apply them to more contexts and realize the need for adjustments in order to preserve some level of reliability (NRC, 2012). Students are envisioned as creating tests for their models to see if they hold true to the nature of the phenomena as they presume (NGSS Lead States, 2013). The key here is that models are useful for being able to predict and relate the phenomena of systems, even while not being a direct representation of reality. Over the course of the school years, students were expected to be able to use more varied and more sophisticated models such as equations, spreadsheets, force-diagrams, etc. (NRC, 2012). By having students not only use existing models, but also develop their own models, it was more likely that they would have to grapple with the assumptions and limitations that were inherent to modeling (NRC, 2012). Developing a metacognitive awareness of the tools that humans use to understand systems provided students with a more true representation of the process of constructing scientific knowledge.

Modeling, as a core practice of most disciplines, also interfaced meaningfully with the other scientific practices. Students could learn to ask scientific questions (*Practice #1*) about others' and their own models (NRC, 2012). Students could design 'stress tests' for their models (*Practice #3*) based on such questions (NRC, 2012). Students could gauge how well models explained data (*Practice #4*). Models could necessitate the use of mathematical or computational representation and manipulation (*Practice #5*). Students could construct explanations (*Practice #6*) for phenomena using models as their reference point of mechanism and relationships (NRC, 2012). There were rich opportunities for students to engage in argument based on evidence (*Practice #7*) as they compared models and considered which models had the greatest merit for certain contexts (NRC, 2012). Finally, the iterative process of refining models was an exemplary opportunity to obtain, evaluate, and communicate scientific concepts (*Practice #8*). When applied intentionally, modeling could be an entrance point to engage in all eight scientific practices.

Modeling in the Broader Literature

Various aspects of model-based inquiry have been explored in the literature since the term was thoroughly established in the 2008 piece by Windschitl et al.: *Beyond the Scientific Method: Model-Based Inquiry as a New Paradigm of Preference for School Science Investigations*. Modeling has the potential to bridge the gap in the content-process dichotomy that has been identified in science education (Passmore et al., 2009; Millar & Driver, 1987). In developing models, students must situate the explanation of content within the scientific process that they undertake (Passmore et al., 2009). For the purposes of framing this capstone project, the planning, enactment, and assessment of MBI will be introduced.

Planning MBI. As opposed to a typical open inquiry where students must identify their own questions and determine their own sources of information, MBI structured the teacher as an agent that chooses the key scientific phenomena that students will explore, with prior knowledge that the phenomenon is undergirded by understandable explanations (Windschitl et al., 2008). This type of support may importantly lower some of the initial cognitive load for novice learners, which was a concern Kirschner et al. raised in their critique of inquiry (2006). Depending on student familiarity with the MBI process, they may benefit from the teacher giving an initial tentative model that takes a stance in relating the features of the system being investigated (Windschitl et al., 2008). Ross and Davidson (2020) laid out how the framework of MBI could also be paired with a framework for orchestrating productive task-based discussions in science called the *Five Practices*, first introduced by Cartier et al. (2013). The teacher could also plan for the lessons by considering in advance the Five Practices in coordination with MBI:

1. *Anticipating*: predicting how students are likely to respond to a task.
2. *Monitoring*: taking purposeful note of what students are doing and saying as they are working on the task.
3. *Selecting*: deciding what information and student ideas will be shared in the whole-class conversation.
4. *Sequencing*: deciding in what order the selected models or ideas should be discussed so students can make sense of the science ideas.

5. *Connecting*: drawing student attention to the connections between the selected student information in a cohesive way that ultimately leads to a complex discussion that helps students further develop the construction of their ideas. (Ross & Davidson, 2020, p. 45)

Only the first practice was able to be prepared in advance, but planning to incorporate the remaining four was a key element to a successful implementation. In the planning process, it was also valuable to understand teacher experiences with novel implementation. Dass et al. (2015) identified five inhibitors of the implementation of MBI as reflected by teacher responses to participating in MBI instruction, one of which related to teacher preparation. The inhibitor was classified as “Adaptation to new inquiry-based pedagogical practices” (p. 1310) and was particularly in reference to the students’ experience of the new approach. Their suggestion for ameliorating this was to prepare students for the structure of MBI by having them first participate in MBI for content knowledge that they are already familiar with before applying it to a novel construct (Dass et al., 2015). Next, the process of enacting MBI during classroom instruction time will be considered.

Enactment of MBI. The actual enactment of MBI is multifaceted and complex. In fact, it has been mentioned that even the process of MBI itself is iterative and over time can yield more successful variations (Neilson et al., 2010). The goal of the MBI process is articulated by Windschitl et al. as “[d]eveloping defensible explanations of the way the natural world works” (2008, p. 15). During the classroom time, the teacher would introduce the phenomenon, give students resources or experiences upon which they can generate an initial model, have students state (possibly multiple) potential relationships between components in a model, use various processes to test the validity of the stated relationships, and then construct an argument in which they attempt to explain the original phenomenon based on how their model fitted the available evidence (Windschitl et al., 2008). Chemistry, in particular, is a subject area that relies heavily on modeling since most of the phenomena being studied are too granular for most instrumentation, let alone the naked eye, to observe.

Throughout this process, the teacher had to be prepared to engage in the last four of the Five Practices of discussions: monitoring, selecting, sequencing, and connecting (Ross & Davidson, 2020). As the teacher monitored, they could identify possible misconceptions, which need not be directly corrected, but which the teacher could use to prompt the students to test the defensibility of that portion of their model with certain data later on (Ross & Davidson, 2020). The teacher could also formally monitor how student models change over time in a unit (Ross & Davidson, 2020). During interspersed discussions with the class, the teacher could select models that have notably changed over time in order to elevate the practice of revising models (Ross & Davidson, 2020). The teacher could use the selected features of different student models during the sequencing phase, where they could call on groups to share certain features from their models so as to build a sequence of understandings that tell the storyline of the phenomena (Ross & Davidson, 2020). Finally, the teacher could have students connect their initial ideas to their updated models to consider how their thinking had changed over time, with the hope that by self-correcting their initial conceptions they could experience true conceptual change (Ross & Davidson, 2020; Anderson, 2008a).

There are a number of different proposed MBI cycles that are worth comparing. Each MBI process that is detailed in prior scholarship is inherently circular and iterative (Dass et al., 2015; Neilson et al., 2010; Windschitl et al., 2008). First, the bidirectional circular process put forward by Windschitl et al. (2008) had students looping around four conversations labeled as: organizing what we know/what we'd like to know, generating hypotheses, seeking evidence, and constructing an argument. The terminal endpoint of this cyclical process was for students to show how they had developed defensible explanations of the way the natural world works (Windschitl et al., 2008). In the road map for MBI presented by Neilson et al., there was a cyclical process of modeling, focused inquiry, and iterations on repeat until the students met rubric criteria for their model (2010). Dass et al. (2015) presented a cyclical process with the following phases: macroscopic exploration, initial model development, peer review of class models, test and apply models in secondary exploration, model revision, and peer review of class models. For each phase of peer review of class models in this cycle, there was the opportunity for implementation of the Five Practices or for a teacher clarification step (Colburn & Clough, 1997; Ross & Davidson, 2020). Since each cycle claims to be rooted in MBI, it is useful to directly relate the crossover between the different models to highlight the consistent features of different conceptualizations of MBI processes. To distill these consistent features, Table 1 presents a side-by-side comparison of three different MBI processes. To gain a deeper understanding of each of the cycles being compared, see the original articles (Dass et al., 2015; Neilson et al., 2010; Windschitl et al., 2008).

Table 1

Shared Iterative Process Between Various MBI Cycles

	Windschitl et al., 2008, p. 15	Neilson et al., 2010, p. 39	Dass et al., 2015, p. 1308
Cycle Description	Bidirectional cycle with entrance and exit points	3 part unidirectiona l cycle	6 part unidirectional cycle
Introduction of Phenomena	*Entrance to cycle:* Setting the broad parameters	*Top*: Modeling	*Step 1*: Macroscopic Exploration
Initial Model Construction	*Top left*: Organizing what we know, what we'd like to know *Top right*: Generating Hypotheses	*Top*: Modeling	*Step 2*: Initial Model Development
Testing of Model	*Bottom right*: Seeking Evidence	*Bottom right*: Focused Inquiry	*Step 4*: Test and apply models in secondary exploration
Model Revision	*Bottom left:* Constructing an Argument *Top left:* Organizing what we know, what we'd like to know	*Bottom left:* Iterations	*Step 5*: Model Revision
Defending Model Based on Evidence	*Bottom left:* Constructing an Argument *Exit to cycle*: Developing defensible explanations of the way the natural world works	*Bottom left*: Iterations	*Steps 3 and 6*: Peer Review of Class Models

Dass et al. (2015) found several themes of inhibitors to the implementation of MBI in the classroom. One such theme was captured in this teacher's response:

> Push back from students initially not being sure what I am asking for. I have honors students and they always want specifics of the requirements. When I leave things open-ended they get nervous but that also provides for an array of answers which is good. (Dass et al., 2015, p. 1310)

To lessen the severity of the inhibitor of student push-back, Dass et al. offered the suggestion of building buy-in for sharing and reflecting in the classroom culture as a whole, since the practice was so essential to MBI (2015). The second inhibitor was typified in one teacher's response: “The issue that I’m having is just trying to find the time with all the other standards that we’ve got to hit in Chemistry" (Dass et al., p. 1310). Their response to this inhibitor was that since the NGSS and AP both detail modeling in their articulation of science standards and objectives that the teacher may proceed without this worry. Finally, they discuss the inhibitor of lacking the material resources or funds to acquire all necessary materials. Their argument in response to this inhibitor was that MBI need not be implemented in only novel investigations, but rather could utilize the materials that had been historically used while reframing the investigation with the MBI lens. While other curricular approaches focus on having air-tight lesson plans and rigid sequencing, MBI relies on the teacher as the primary resource and guide. The teacher must be found engaging dialogically with the students to guide them in refining their thinking over time and to press them to argue from evidence (Ross & Davidson, 2020). Windschitl et al. (2008) detailed example questions to implement throughout the conversations within a MBI cycle. These will be discussed in greater detail in Chapter Three. It is also worth noting that MBI requires teachers to have a deeper understanding of the disciplinary concepts in order to guide students through the conversations (Windschitl et al., 2008). This could lead to issues in scaling MBI as an effective teaching strategy when not all science teachers in the US are teaching within an area of their expertise. Conversely, Dass et al. found that teachers who implemented MBI *gained* further content knowledge through implementing MBI, so there may be an argument for MBI being scalable in that over time the teachers may gain a greater familiarity and skill with the disciplinary content (2015). Although MBI need not be inherently experimental, there may also need to be teacher demonstrations and student safety training in regards to experiments and using instrumentation in their exploratory process (Neilson et al., 2010).

Assessment in MBI. MBI facilitates rich opportunities for both informal and formal formative assessments and authentic summative assessments. Since MBI is inherently an iterative process, there is nearly constant formative assessment at play while the teacher engages in monitoring, selecting, sequencing and connecting student ideas (Ross & Davidson, 2020). A teacher-perceived catalyst making MBI worth implementing was how it gave ample opportunities for both teacher and students to recognize and correct misconceptions along the way (Dass et al., 2015). In regards to formal formative and summative assessments, there are strengths and considerations. Windschitl et al. did note that teachers implementing MBI will need to have a “broader repertoire of assessment practices” in order to accomplish such ambitious pedagogy (2008, p. 23). Since the modeling products are going to differ by student and by group, it was important for the teacher to develop their own exemplar from which to generate rubric language for the students (Ross & Davidson, 2020). Explicit rubric language could reassure teachers of objective features to assess work by while still keeping the representation open to be molded by the students' creativity and process (Dass et al., 2015; Ross & Davidson, 2020). See Neilson et al. (2010) for a physics MBI rubric on the concept of buoyancy force.

Summary of Limitations of MBI. Since MBI is still in its early stages of development, there is little empirical exploration of its effects relative to traditional teacher methods or generic inquiry. Ogan-Bekiroğlu and Arslan (2014) found that among pre-service teachers in Turkey that implemented either inquiry-based or model-based inquiry there were not any differences in the students' process skills and conceptual knowledge. While this study was a true-experiment with even the merits of randomization leveraged, its sample size was only 26 total teachers, making it necessary to gauge if this result is reliable over multiple additional settings (Ogan-Bekiroğlu & Arslan, 2014). As it stands, a conservative consideration would be to assume that MBI is not any more effective than inquiry-based instruction until there is data suggesting otherwise. It is reassuring, however, that there also was no marked decline in student results relative to inquiry instruction, making it reasonable to pursue further study. Conducting further high quality, large scale, randomized studies is imperative, especially if MBI is to be considered as a scalable reform pedagogy. Lastly, since MBI is an ambitious pedagogy that synthesizes numerous nontraditional teaching practices, effective implementation will depend greatly on the training and skill of teachers (Windschitl et al., 2008). It is obvious that there will be variation in teacher implementation, so measuring modulators of effect for MBI will also be useful to research further.

Summary of Strengths of MBI. Model Based Inquiry is a promising form of inquiry due to its many merits. MBI may be a corpus callosum for the content-process dichotomy that has plagued science education (Passmore et al., 2009; Millar & Driver, 1987). It readily shifts the classroom experience from being teacher-centered to being student-centered (Dass et al., 2015). In conjunction, students become active learners rather than passive learners (Dass et al., 2015). There is also a more convincing parallel with scientific investigation since modeling correctly represents the nature of scientific knowledge as being testable, revisable, explanatory, conjectural, and generative (Neilson et al., 2010; Windschitl et al., 2008). The collaboration, creativity, and flexibility embedded within the MBI process –paired with responsive teacher guidance–sets the stage well for frequent teacher feedback and conversation (Dass et al., 2015). These aspects lend themselves to quality differentiation in the classroom. Overall, it is reasonable to consider MBI as a possible successor to the inquiry-based instruction paradigm in science education, but it is the burden of education researchers to generate empirical evidence to support, refute, or refine this consideration going forward.

Summary of Modeling in Science Education

As one of the eight core Science and Engineering Practices (SEPs) presented by the *Framework* (NRC, 2012), modeling has been illuminated as a rich entrance point for student science learning. In addition, it has been identified as being particularly potent for bridging the process-content dichotomy (Passmore et al., 2009). The process of representing tentative mechanisms and relationships for the parts of a system is useful for scientifically investigating and understanding phenomena but also for gaining an appreciation for the way that scientific knowledge is generated, tested, and refined (NRC, 2012; Windschitl et al., 2008). With the introduction of model-based inquiry there are now detailed suggestions for planning, conducting, and assessing the practice of modeling in the science classroom (Dass et al., 2015; Neilson et al., 2010; Ross & Davidson, 2020; Windschitl et al., 2008). While the literature about MBI is growing, there has yet to be a large-scale seminal empirical study for MBI like that conducted by Geier et al. (2008) for inquiry-based instruction.

The capstone project I have designed utilizes the framework of MBI to create a protocol for students investigating non-experimental chemistry concepts while engaging in scientific practices. The data sets are designed to serve as the exposure to human science conventions or models that students then mine for patterns and build their own models for in order to elucidate the chemistry concepts presented therein. The protocol then details the role of the teacher in the guidance of the process as students wrestle with how to make their models best fit the available data or information. Chapter Three will give a detailed account of how the data sets will be designed and what the protocol will suggest as the two component parts of this curricular capstone project.

CHAPTER THREE

Project Description

Introduction

While the discussion around inquiry in science education has enjoyed predominance for decades, teachers in the United States continue to feel tension in the tradeoffs they perceive in incorporating inquiry approaches (Anderson, 2008b). There is concern about the time needed for inquiry investigations, whether inquiry is realistic in context, and the unfamiliarity of the roles for teachers and students (Anderson, 2008b). There was also a formal critique of science education being stuck in a predictable pendulum swing of a content-process dichotomy, with the modern paradigm overemphasizing scientific process (Millar & Driver, 1987). Model Based Inquiry (MBI) was identified as a powerful iteration of inquiry-based science education that represented authentic scientific practice and pushed students to deeply understand disciplinary content (Passmore et al., 2009; Windschitl et al., 2008). The curriculum design framework of Understanding by Design (UbD) was also utilized to address this tension (Wiggins & McTighe, 2011). The UbD framework establishes short-term acquisition and meaning goals that connect to the disciplinary content knowledge while simultaneously pursuing long-term transfer goals such as student skill in applying scientific practices (Wiggins & McTighe, 2011). I created this project to address the felt tension of teachers wanting to shift to a better balance of content and process, as exemplified in the question: *What kind of curricular resource effectively synthesizes the parallel goals of engaging students in scientific practice and understanding disciplinary content?* This capstone project was designed to give high school chemistry teachers a disciplinarily contextualized and scientific-practice-infused curricular resource that provides a didactic synthesis to the content-process dichotomy. This capstone project presented a MBI-inspired protocol paired with strategically designed data sets reflecting non-experimental chemistry content.

As a science curricular resource, this capstone project addressed a number of important considerations. First, the project components themselves were fully described and rationalized. This included the data sets, the MBI-inspired protocol, the UbD rationale and the supporting components. Next, the project description indicated the setting, audience, and timeline of this project. Lastly, the nature of how this curricular resource could be assessed for effectiveness in the classroom was proposed, even though it was beyond the scope of this project to carry out these suggestions. Introducing the project components first was the best way to contextualize the project.

Project Components

The curricular resource of this project consisted of two main artifacts: the MBI-inspired protocol and the data sets designed to be paired with the protocol. The non-experimental chemistry content-based data sets each supported performance expectations in specific NGSS disciplinary content standards and served as the selected trailhead from which the students developed their models to represent their inductive thinking regarding the data sets. The MBI-inspired protocol presented a structure for how the teacher could sequence classroom activities in a way that supportively guided students in their development of a model. Combined, these two curricular resources gave specific form to the goal of engaging students in scientific practice and understanding disciplinary content.

Data Sets

The term *data* tends to elicit a particular meaning when referenced but is used in a flexible way in this project. Data sets in this project were considered to be any graphical, pictorial, mathematical, linguistic, physical, simulational, or model-based set of information that could be *mined* for patterns and relationships. A set of values in a table could certainly be considered data but so could a set of examples and non-examples demonstrating the conventions of writing chemical formulas. Each data set was designed to hold within it the pattern for at least one, but often multiple, conventions and concepts in chemistry. These patterns and relationships were the substance of the disciplinary content that was a desired learning outcome for the students. Some of the patterns and relationships were directly related to natural phenomena while others were related to important human models and conventions in chemistry.

As an example, consider developing a data set demonstrating the use of significant figures in precisely reporting measurements. Digits in a number are considered significant figures based on a number of key rules. One could develop a data set that showed examples of numbers accurately rounded to one significant figure, numbers rounded to two significant figures, and numbers rounded to three significant figures. By looking at the various examples, students could begin to develop models for representing which types of digits are considered significant and not. Starting with observing this data set could then initiate them into the journey through the MBI-inspired protocol until they had a data-fitted model useful for describing or explaining the system being explored.

MBI-Inspired Protocol

The MBI-inspired protocol was developed by fusing elements of previously proposed MBI cycles. As explored in the literature review, there were a few key components in each MBI cycle. These five elements were presented in Chapter Two in Table 1 but are repeated here:

1. Introduction of Phenomena
2. Initial Model Construction
3. Testing of Model
4. Model Revision
5. Defending Model Based on Evidence

A new MBI-inspired protocol was generated where each of these five elements were present. The protocol itself had its own model to demonstrate how these five features were represented. A phenomenon storyline and essential questions were established for the content exploration. Model development was the central activity that was visited and revisited after various input phases. Input phases included data exposures, peer review of models, and focused-testing of models. The exit product of the process was to have generated a data-fitted model that was useful for describing and explaining the system explored.

The phenomenon storyline and essential question were elements that introduced the phenomenon that the exploration was situated within. These two elements presented a greater *why?* for the exploration and a focused curiosity for students to consider as they were informed in their model development process. Others have built out curricular resources for phenomena in earth science and physics (Neilson et al., 2010; Windschitl et al., 2008). There is also one particularly insightful chemistry MBI example presented by

Ross & Davidson (2020). Nonetheless, there remained a dearth of examples of model-based inquiry for high school level chemistry, particularly for non-experimental chemistry content.

While most chemistry knowledge had been discovered and validated through experimentation, I used the term *non-experimental* to refer to chemistry concepts that high school students generally were not able to learn from their own experimentation due to the limitations of high school laboratory equipment. For example, most high school labs were not equipped to reproduce Rutheford's gold-foil experiment to discover the dense, positively charged nucleus of an atom. Therefore, the characteristics of subatomic particles in an atom were considered *non-experimental* chemistry content. Since all of chemistry, by nature, relies heavily on models to represent the unobservable atomic level, MBI offered a particularly relevant framework for approaching even the non-experimental chemistry topics that were covered in high school chemistry classrooms. The process of MBI still needed to be initiated from a relevant context and curiosity which was articulated in the phenomenon storyline and essential question.

After the phenomenon storyline and essential question set the stage for the exploration, the next step was the initial data exposure. This was where the data sets described above were first used. The data exposure could be *mining* existing models in chemistry for the rationale of their features, or from *mining* data sets that represented conventions used in chemistry for their rules and nuances, as described previously.

The core activity of the process was the students' development of their own models for the system being explored. Students needed some chosen medium to create their model with and a group to work with. They could use large whiteboards, Google slides, Google docs, or some other medium that all students in their group could contribute to. They could make choices about what and how to represent the features of the system explored and the relationships between features therein.

The MBI-inspired protocol also included a phase for peer review of models. This phase was important as it gave students greater familiarity with the features and relationships of the system as they considered alternative ways to represent the system in a model. It was also meaningful because students got the opportunity to practice communicating about and assessing the power of models to describe and explain a system.

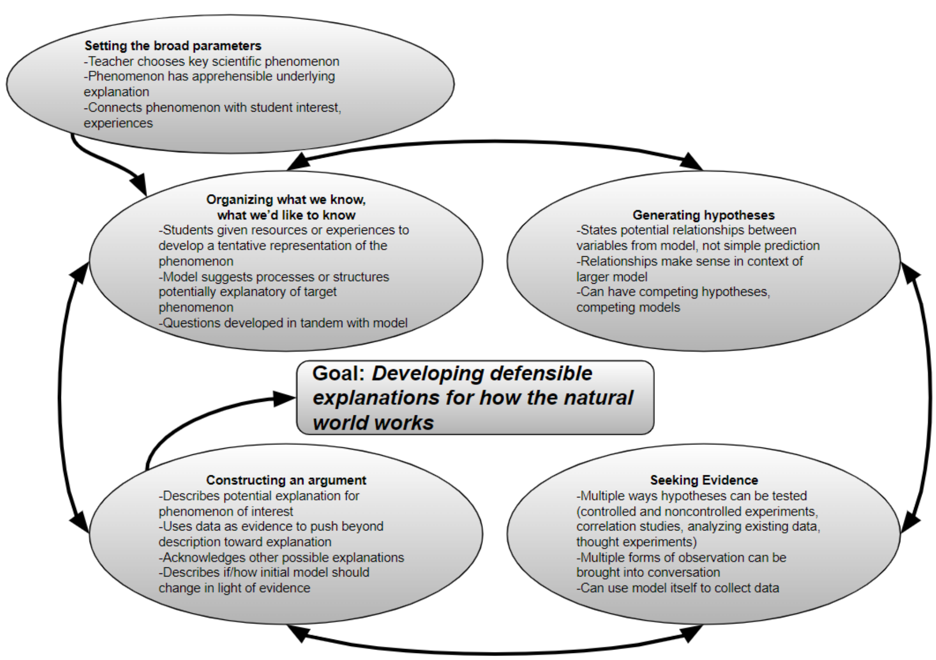
The last input phase was a focused-testing of models phase. Here, students had their models tested by strategically chosen prompts and questions that the teacher prepared. Each prompt or question demanded students to commit to their definitions and explanations of the features and relationships represented in their model. Here, they saw if their models held up to established understandings of the concepts being explored, which the teacher was aware of but they likely were not. The teacher could ask a focused question about the features in the model, such as: “what particles are present in the middle of the atom?” Students then used their model to state their prediction or claim. The teacher could also probe about a relationship between features with a question like: “how are the number of electrons and protons related?” The teacher could also ask students to use their model to predictively generate something, such as: “create a Bohr Model for the atom of Neon based on your current model.” Through a series of prompts and questions the nature of the features and relationships in their model was systematically tested for reliability, generalizability, and limitations.

Finally, the protocol needed some sort of exit product that held students accountable to demonstrating their progress towards the goals of the exploration. This exit product was a demonstration of their ability to use a data-fitted model to describe and explain the system being explored. This was where a formal formative or summative assessment could be used.

Each phase of the MBI-inspired protocol included a list of questions to press students with as they explored the content and developed their models. These pressing questions were inspired by the four conversations model presented by Windshitl et al. (2008). Figure 1 is a rendition of this MBI cycle provided by Windschitl et al. (2008). The four conversations were pursued iteratively and were the bridge between students experiencing a phenomenon and them being able to develop a defensible explanation of how that phenomenon works (Windschitl et al., 2008).

Figure 1

The Four Interrelated Conversations Supporting Model-Based Inquiry



Note. This model was presented by Windschitl et al. in 2008, depicting the components and possible sequences of the MBI process. Adapted from "Beyond the Scientific Method: Model-based Inquiry as a New Paradigm of Preference for School Science," by M. Windschitl, J. Thompson, and M. Braaten, 2008, *Science Education 92* (5): 941–967. Copyright 2008 by Wiley Periodicals, Inc.

Windschitl et al. provided a set of questions to be explored in each of the four conversations but cautioned that these conversations must not be considered to be linearly pursued, but rather that there was constant revisiting of earlier conversations as students tested, developed, and reasoned with their model (Windschitl et al., 2008). This caution was represented in the bidirectionality of the arrows in Figure 1. Each conversation phase in the cycle could be guided and supported by the teacher pressing the students to consider various questions and considerations as they inductively reasoned their way through the data exposures and model testing. Table 2 presents such artifacts of communication that might be implemented by a teacher when they are facilitating conversations throughout the MBI-inspired protocol. This set of questions was representative of the questions that Windschitl et al. offered in their 2008 article.

Table 2

Artifacts of Communication in the Four Conversations

Conversation	**Artifacts of Communication**
1: Organizing What We Know and What We Want to Know	1. What do we already think we know about this situation, process, or event? 2. Which of the elements of what we know are based on observation? Inference? Other sources of information? 3. Could there be more than one way to represent this situation, event, or process? A conceptual diagram? Pictorial drawing? 4. Is our model purely *descriptive* of the situation, process, or

	event, or does it have parts that try to *explain* what is happening? 5. What are we leaving out of our model and why? 6. What questions does this model help us ask? 7. What additional information do we need to improve our initial model before asking our final inquiry questions? 8. How can our questions be framed so that they can be answered by collecting and analyzing data?
2: Generating Testable Hypotheses	1. What aspects of our models do we want to test? 2. Are there competing hypotheses that could both explain what we are observing? 3. When we look at our tentative model and consider the question we want to ask, what would our models predict? 4. How can we test our models in a way that generates better descriptions of how this phenomenon happens? 5. Can we test our models in a way that helps us understand some process or thing that explains why the models work the way they are predicted to?
3: Seeking Evidence	1. What kind of data would help us test the ideas in our model? 2. How can we operationalize our variables in ways that will allow us to record unambiguous measurements? 3. What will it mean to collect data “systematically”? 4. To test our hypothesis, should we observe the phenomenon as it is or actively manipulate some variables while controlling others? 5. Can we use a model itself to collect data from (a tabletop watershed model, a bell jar with rubber diaphragm as a model of the lungs, a computer simulation of population genetics), and will that give us valid data? 6. When we analyze our data, will we compare groups? Look for correlations between variables? Seek other kinds of patterns and trends? 7. What forms of representation (tables, graphs, charts, diagrams, etc.) are most appropriate for the type of data we will collect? 8. How should we handle data that are unexpected or run counter to our hypothesis?
4: Constructing an Argument	1. Was what our original model predicted consistent with the data we collected? 2. How can we go beyond arguing for a simple “cause-and-effect” relationship? For a noncausal correlation between variables?

	3. Are there intermediate or confounding variables we have overlooked? 4. Did we control or at least account for variables that could have influenced the outcomes of the study? 5. How do the data help us infer about theoretical (unobservable) events in our model? 6. How consistent and coherent is our final explanation for the phenomenon of interest? 7. Are there other possible explanations for the data, and if so, how strong is the evidence for these alternatives? 8. What did our model fail to predict and why? 9. Should our model change in light of the evidence? 10. How can our model be applied to other phenomena?

Note. Compiled from "Beyond the Scientific Method: Model-based Inquiry as a New Paradigm of Preference for School Science," by M. Windschitl, J. Thompson, and M. Braaten, 2008, *Science Education 92* (5): 956–960. Copyright 2008 by Wiley Periodicals, Inc.

Since the curricular resource I developed was focused on utilizing MBI for the non-experimental chemistry content that students were expected to learn, it was helpful for me to adapt the artifacts of communication to better correspond to an MBI investigation of chemistry data sets that the students do not generate from experimentation. For example, an artifact of communication that was not relevant in a non-experimental MBI process was: "How can we operationalize our variables in ways that will allow us to record unambiguous measurements?" (Windschitl et al., 2008, p. 958). This artifact was removed in my compilation of artifacts of communication since it applied to MBI processes that involved direct experimentation by students. There were also a number of questions that were worded in ways that I wanted to adapt to be more precise, concise, accessible, or open-ended. For example, the following artifact was lengthy and structured as a *yes/no* sort of question: "When we analyze our data, will we compare groups? Look for correlations between variables? Seek other kinds of patterns and trends?" (Windschitl et al., 2008, p. 959). In my compilation, I adapted this question to be more concise and open-ended: "When we analyze the data, how will we look for patterns and trends?" I also adjusted the language to directly connect to the data sets, as opposed to referencing an undefined "phenomena." Finally, to give teachers a sense of what these questions are used for in the protocol, I retitled them "Questions to Press With" instead of "Artifacts of Communication" and changed them to second-person point of view instead of first-person. The full list of adapted "Questions to Press With" is found in Table 3.

Table 3

"Questions to Press With" for MBI-Inspired Protocol

Protocol Phase	**Questions to Press With**
Data Exposure (initial or additional)	1. What questions can you pose about the data you are looking at? 2. What do you think you already know about the data set you are looking at? 3. What are all the features of the data set that you notice? 4. What patterns and relationships might be important?
Developing Model of System	*Initial Model Development:* 1. How do you want to represent what you initially think? As a graph, picture, mathematical expression, arrangement of words, a physical construction, simulation, or as an adaptation of an existing model? 2. What are you leaving out of our initial models and why? 3. What questions do your initial models bring up? 4. What questions can you ask about your model? 5. What patterns and connections are you considering in your model? 6. What predictions could someone make from your model? 7. How well does the original data set confirm the ideas in your

	model? 8. What kind of additional data or information would help you further test the ideas in your models? 9. Which alternative ideas are worth simultaneously considering as you are exposed to more data? *Further Development of Model* 10. How should your model be strengthened after seeing the representations in other groups' models? 11. How do you choose which feedback from your peers to act on and what to ignore? 12. How should your model change in light of the further data exposure? 13. How should your model change after being challenged by the focused-testing prompts? 14. What additional features or variables might still be overlooked or unexplored in your model? 15. How is your final model consistent and coherent for explaining the totality of input and information you were exposed to?
Peer Review of Class Models	1. What competing models might also explain what is represented in the data set? 2. What features and relationships did others include that you did not include in your model?
Focused Testing of Models	1. How did the predictions of your initial model hold up with the additional data? 2. What did your model fail to predict and why? 3. How should you handle results that were unexpected or do not support your models? 4. How should your model change in light of the evidence? 5. How will you represent your findings in your models? 6. What other contexts may your model be applied to?

Note. Some of these questions were inspired, adapted, or directly chosen from "Beyond the Scientific Method: Model-based Inquiry as a New Paradigm of Preference for School Science," by M. Windschitl, J. Thompson, and M. Braaten, 2008, *Science Education 92* (5): 956–960. Copyright 2008 by Wiley Periodicals, Inc.

Understanding by Design Rationale

An important consideration in the development of this curricular resource was to align it with relevant and meaningful goals. To this end, the framework of UbD was used to clarify the short-term disciplinary content meaning and acquisition goals and long-term scientific practice transfer goals that the project met (Wiggins and McTighe, 2011). For the data sets, the NGSS language was used as a starting point to identify meaningful learning targets in the domains of acquisition (what students will know and be skilled at) and meaning (what students will understand and be considering). In this way, the data sets accomplished short-term goals for disciplinary concepts. The MBI-inspired protocol was designed with learning targets in the domain of transfer (what students can use their learning to do beyond the specific content covered). This design choice helped tighten adherence to long-term goals for students developing confidence in scientific practices.

Supporting Components

Since one goal of this project was to give high school science teachers in general a vision of how they could develop similar data sets for other content, the first two data sets presented in the protocol also included supporting components. The supporting components are described and rationalized in Table 4. The third column designated which curricular artifact the supporting component was found within. Some components were included in the data set only because the data set components were meant to represent all the materials that a teacher would prepare for their content. Other supporting components were only found in the protocol as it was meant to detail the process that the teacher went through with the students, which included phases that the teacher could not prepare (e.g. the students' models, peer review of models, and conversations).

The combination of the data sets and the supporting components was essential for a teacher to envision the initiation and end-goal of the modeling process while the MBI-inspired protocol itself provided greater detail as to how the input phases and model development phases could be sequenced and related. With the project components described in detail, it is now worth establishing the larger context in which the project was intended to be implemented, namely Minnesota public and charter schools.

Table 4

Descriptions of Supporting Components

Supporting Component	**Description**	**Curricular Artifact**
Non-Content Example	A non-content example will be included so that teachers of any subject can see how the protocol can be used for a wide range of content. It will also be a more accessible example for teachers unfamiliar with chemistry content. Similarly, it is a useful introductory exploration for students to become familiar with the process of model development.	Protocol and Data Set
Alignment with NGSS	The relevant NGSS components will be explicitly presented to demonstrate the concordance between this protocol and the NGSS.	Data Set
Alignment with Understanding by Design (UbD)	UbD elements for the exploration will be explicitly articulated. This will include consideration of an essential question, short-term acquisition and meaning goals, and long-term transfer goals.	Data Set
Notable Features of Data Set	A list of notable features of the data set will be included to “pull back the curtain” so that teachers know what rationale is used in strategically designing the data sets.	Protocol
Student Example Initial Models	The example student initial model will give teachers a possible construction of a model, similar in substance to the ones students may make during the modeling process. Knowing what a model could look like is vital as the teacher seeks to push their students to refine their models in the process. The example model also gives the teachers context as to what features, patterns and relationships they could ask students about with the Questions to Press With.	Protocol

Data Exposure (additional)	Additional relevant data provides more examples of the possible relationships and patterns, introducing greater complexity, nuances and exceptions. The additional relevant data is meant to give students the opportunity to test their models and see if the predictions of their model are reliable for the context.	Protocol and Data Set
Example of Peer Review of Models	Every peer review of class models will look different because of the different students engaging in the process, but an example of possible peer review feedback is given to demonstrate the type of peer feedback students will be encouraged to provide.	Protocol
Exit Product	The assessment is the place where they demonstrate their inductive thinking regarding the totality of the data or information they looked at. This is where students are held accountable for being able to use their data-fitted model to describe and explain the system they explored.	Protocol and Data Set

Setting

This curricular resource was intended to be implemented in Minnesota public and charter schools aligning with the NGSS. In the state of Minnesota, high school students were required to complete three credits in science, including either chemistry or physics (MDE, n.d.b). Because of this, high school chemistry remained a predominant science course offered in most public and charter schools in Minnesota. While not all public and charter schools had fully transitioned to the NGSS to date, many were in the process of switching. The Minnesota Department of Education (MDE) set their intention to modify the Science Minnesota Comprehensive Assessment (MCA) according to their NGSS-aligned 2019 standards by the 2024-25 school year and therefore many schools were phasing in the 2019 standards ahead of this update (MDE, n.d.a).

The students in Minnesota public and charter schools were diverse along multiple identifiers. The following statistics were drawn from the MDE public schools enrollment data for the 2019-2020 school year, which was the most recent publicly available data (MDE, 2021). Students who identified as Black, Hispanic, Asian, Native American, and multiracial collectively accounted for over 35% of public and charter school students in Minnesota (MDE, 2021). Students receiving special education services accounted for over 16% of students in Minnesota. A meaningful portion of students (~8%) were English Language Learners (ELLs). Finally, over 35% of students in Minnesota were eligible for free & reduced lunch, which was an approximate surrogate for lower socioeconomic status.

Because of the diverse demographics of Minnesota schools, the successful implementation of any NGSS-aligned curriculum could be assumed to be reliant on it being culturally inclusive, differentiated, supportive of language development, and poverty-informed. The NGSS noted how their framework was meant to make science learning accessible to all students. Citing Lee and Buxton (2010), the NGSS Lead States referenced three areas of importance in equitably teaching students from diverse and underserved backgrounds:

> (1) value and respect the experiences that all students bring from their backgrounds (e.g., homes or communities), (2) articulate students' background knowledge (e.g., cultural or linguistic knowledge) with disciplinary knowledge, and (3) offer sufficient school resources to support student learning.
> (Volume 2: Appendixes, p. 30, 2013)

These three supports were important features of successful implementation of NGSS-aligned curriculum, considering the diverse context of Minnesota schools. In addition to describing the features of the student population in Minnesota, it was also important to indicate the target audience that this project was meant to be relevant for.

Audience

The curricular resource presented in this capstone was intended to be used by preservice and current high school chemistry teachers in Minnesota public and charter schools. This project was relevant and timely because most high school chemistry teachers across Minnesota needed to adjust their curriculum over the next two years to meet the expectations of the 2019 Minnesota science standards and prepare their students for the 4th iteration of the MCA. Change is hard on human cognition. An analogy from chemistry is quite appropriate here. During a chemical reaction, the reacting molecules often must reach a *transition state* that requires higher energy before they can reach a more stable end-product. There may be a temptation for chemistry teachers to avoid such a *transition state* and instead justify their existing curriculum as already being NGSS-aligned, even if that justification is dubious. An exemplary curricular resource could be a particularly helpful catalyst in this transition because teachers would have access to a contextualized, well-justified, and inspiring reference point for NGSS-aligned chemistry lessons to consider as an alternative option to their current practice. The word catalyst was used very intentionally above. In chemistry, catalysts, by definition, *lower the necessary energy of the transition state*, making it less strenuous for the molecule to reach the more stable end-state. For a realistic transition to the new standards, curricular catalysts like the one presented here could be of much help.

Timeline

This capstone project was completed over the summer of 2023, in hopes that it could be distributed to chemistry-teaching colleagues throughout Minnesota for use in the 2023-2024 school year. Having this curricular resource available for chemistry teachers to try in the 2023-2024 school year was timely and helpful since many of them needed to be phasing in NGSS-aligned curriculum ahead of the 2024-2025 school year when the MDE officially updates the MCA to assess the new 2019 Minnesota standards.

Assessment

This capstone project was founded on educational research in inquiry and particularly on the more recent iteration of inquiry-based instruction, MBI. To understand the nature of the effect size of implementing this curricular resource, tests needed to be done at various scales. It was not within the scope of this project to plan and conduct these necessary tests because the project was designed to only present a curricular resource justified on existing scholarship and evidence. Nonetheless, in Chapter Four, I presented a series of possible research pathways to establish an empirical evidence base for consideration of using this project.

Summary

This capstone project was designed to present a useful curricular resource that high school chemistry teachers in Minnesota could implement as they transitioned to the Minnesota 2019 NGSS-aligned standards over the following few years after this publication. The curricular resource components included an MBI-inspired protocol and corresponding data sets with supporting components to aid the teacher new to the process. The project was completed over the summer of 2023 and was ready to be disseminated to current Minnesota high school chemistry teachers for use in the 2023-2024 school year. While these aforementioned details summarized the scope of the project, it was hoped that an empirical study of the nature and effect size of this curricular resource would be conducted. The project itself is presented next, followed by Chapter Four, a reflection on the process and learnings in developing the project.

CHAPTER FOUR

Project Reflection

Introduction

The project that I developed sought to answer the question: *What kind of curricular resources effectively synthesize the parallel goals of engaging students in scientific practice and understanding disciplinary content?* The resulting project was not a typical unit of curriculum, but rather a rationalized curricular model of how to utilize strategically designed data sets with a model-based inquiry (MBI) inspired protocol to coordinate the parallel goals of students learning disciplinary content and scientific practices. The development of this curricular model proved to demand much metacognitive struggle and reflection.

The metacognitive landscape traversed in developing this curricular model was composed of a number of elements. The primary and most consequential element was my thinking around the implications and limitations of my project. The next major reflection surrounded my scope and understanding of educational research itself, particularly my awakening to how difficult it is to study educational interventions. As an extension of this reflection, I also propose a possible future research path. The final portion of reflection was my struggle in situating myself and this project within the larger scholarship regarding MBI.

Project Implications and Limitations

This project has important implications for current Minnesota science teachers but also considerable limitations to disclose. This project is a curricular resource presented at a uniquely pointed transition in what science standards are being used in Minnesota. The curricular model reflects key shifts in science education proposed in the Next Generation Science Standards. Furthermore, the curricular model was designed with consideration of the barriers that keep many teachers at an arm's length from inquiry focused learning in the classroom. Because of its various merits, I plan to use and disseminate this resource so that it can be incorporated by myself and other teachers. The limitations of this project revolve largely around its inability to claim empirical validation until further research is conducted. Other limitations include its scope and format. I will start by detailing the implications of the project and then discuss the limitations.

Timeliness

This project is particularly timely for current high school science teachers and current pre-service teachers in Minnesota. Because the Science Minnesota Comprehensive Assessment (MCA) will reflect the NGSS-aligned 2019 Minnesota standards by the 2024-2025 school year, this upcoming school year, 2023-2024, is the last year for science teachers in Minnesota to adopt NGSS-aligned curricular approaches ahead of the changes in testing (MDE, n.d.a). Therefore, as a NGSS-aligned curricular resource, my project is a relevant tool for teachers beginning to imagine and implement approaches that incorporate scientific practices alongside the disciplinary content. Furthermore, this curricular model mirrors the priorities set by the NGSS where content is not covered as broadly, but rather more deeply as it is more meaningfully paired with the learning of scientific practices and cross-cutting concepts in science (NGSS Lead States, 2013). This curricular model emphasizes the role of exploring phenomena, identifying patterns, analyzing data, developing and assessing models, communicating scientifically, and describing and explaining systems using evidence.

Addressing Barriers to Implementing MBI

From my own experience teaching, and from my review of the literature, I became aware of the hesitations and perceived barriers that teachers have when it comes to implementing model-based inquiry approaches in the classroom (Dass et al., 2015). This project was designed to ameliorate a number of those barriers. Some real and perceived barriers include discomfort for teachers and students in adapting to a new pedagogical approach, time management in covering all the required standards, and access to necessary materials (Dass et al., 2015).

To help ease the teacher and the students into greater comfort with using the new pedagogical approach of the MBI-inspired protocol, both a non-content example and a content-based example are presented side-by-side. The example of exploring the model of maps is one that is readily accessible to most teachers and students, so there is room for the learning to be focused on the process rather than both process and content. Furthermore, seeing the content example next to the non-content example allows teachers to see the consistency in the structure of the protocol for different data sets and topics.

Additionally, the protocol is designed with a gradual release format, allowing teachers to have significant support initially imagining and rationalizing the process, but then asking more of the teacher in imagining subsequent processes. First, the protocol is presented with two examples side by side, giving teachers the opportunity to clearly see the parallel structure across topics. Paired with this, teachers are given a secondary curricular model that gives a specific sequence to the process so that the choices they are asked to make initially are minimized. In the side by side examples, each phase of the protocol is rationalized and visualized by accompanying examples. After the dual examples are shared, the next step in the gradual release is the presentation of an additional protocol cycle with just the components that a teacher would prepare, provided along with some guiding questions considered in developing each element. Finally, a blank version of the data set planning tool is given, with the guiding questions remaining. By designing the curricular resource with gradual release in mind, the teacher's journey to utilizing the resource is scaffolded, which minimizes the stress of trying a new pedagogical approach.

The perceived barrier about MBI taking too much class time away from covering the standards may have been defensible before the 2019 Minnesota standards, but now that standards are being aligned with the NGSS, it is paramount that teachers actually prioritize approaches that incorporate scientific practices alongside the narrower set of disciplinary content topics. Assuming that the new wave of the Science MCA demonstrates integrity to the performance expectations laid out in the NGSS, teachers can no longer afford to avoid teaching scientific practices explicitly alongside disciplinary content.

Finally, the concern around access to materials is largely irrelevant in relation to this curricular model. While other scholarship addresses how to approach experimentally derived chemistry knowledge with a more congruent inquiry lens and therefore may require particular equipment and materials, this project explicitly aims at helping science teachers imagine even non-experimentally based content knowledge as a trailhead for students to use scientific practices to arrive at that disciplinary content knowledge.

By addressing important real and perceived barriers to teachers' implementation of MBI, this project is offered as a high quality resource for current and pre-service teachers in Minnesota. By being thoughtfully designed, scaffolded, and focused in scope, this curricular resource is made accessible and useful for teachers who have kept inquiry processes at arm's length.

Intentions for the Project

I will be seeking to use and disseminate this resource in numerous ways. Personally, I will be using this curricular resource in my own curriculum as a chemistry teacher this upcoming school year at a 6-12 private school in the upper midwest. I will also be sharing this resource with colleagues at my current and previous institutions as they grapple with the landscape of the transition of standards. Additionally, I will be submitting my curricular resource for consideration on the Model Based Inquiry resource hub website that is run by two professors at Northern Arizona University and Neag School of Education, Dr. Ron Gray and Dr. Todd Campbell (Gray & Campbell, n.d.).

Limitations

The limitations of this project lie in its scope, format, and particularly in its empirical validity. The curricular model presented initiates the journey toward meeting the standards for atomic structure (HS-PS1-1) and organization of periodic table (HS-PS1-2) (NGSS Lead States, 2013) but does not give the scope of a full unit. This choice is intentional because the curricular resource is meant as one tool to consider for curriculum development, not as a single comprehensive tool in a curricular approach. Knowing that some chemistry knowledge is best informed by experimental exploration, this project is meant to inspire teachers to also consider how scientific practices like modeling can be infused throughout non-experimental content as well.

The second limitation is the format of the project. As a curricular resource document with a gradual release design, it can be helpful and scaffolded for teachers willing to try it as a new approach. Nonetheless, the format of a document is limited in its ability to change or support teacher practice. Additional support for teachers such as professional development and ongoing mentorship may be important additional formats to consider. For example, in the research conducted by Dass et al. (2015), teachers benefited from the structure of ongoing meetings throughout the school year to discuss using modeling in the science classroom.

The most consequential limitation of this project is its inherent lack of empirical validation. While firmly rooted in empirical research on inquiry and scholarship around MBI, this use of this project itself has not undergone any sort of rigorous empirical validation. The extent to which the effect of the project components have been observed are purely anecdotal to my own teaching practice over the last three years, and the actual final curricular model has only been used to guide instruction in one cohort of students in a midwestern public high school summer academy. These settings in which the project elements have been explored have hinted at its power to engage students in scientific practices through the central activity of model development, yet these anecdotal accounts are not only uncontrolled but also highly vulnerable to sample bias, selection bias, reporting bias, and various forms of confounding. For these reasons, I claim no *policy* implications based on my project until further empirical validation of the use of this curricular resource has been conducted. Even without claiming my project should change education policy, I still offer this project as a high quality curricular resource, rooted in external research and scholarship, for individual teachers and schools to consider as they seek to adjust their practice to better align with NGSS standards in the coming years.

In summary, this project is a curricular resource rooted in research and scholarship with a number of key implications in the immediate future for Minnesota science teachers. The project also had important limitations and cannot be used to inform policy at this point, yet it is still a useful, research-based resource to current and preservice science teachers. As a tool aligned with NGSS standards, it can be a valuable curricular resource for Minnesota science teachers shifting their practice with the shift in standards and the MCAs. The project is designed to minimize the barriers for teachers new to teaching scientific practices and inquiry. I will be utilizing my project in my own teaching practice, disseminating it for use by colleagues in other educational institutions, and seeking to have it available online to educators. My second major area of reflection from developing this project is how it has piqued my interest in exploring and analyzing educational research itself.

Reflecting on Educational Research

I was surprised by how significantly my perspective of educational research was impacted by the process of researching for and developing my project. I experienced cognitive dissonance in designing a project without empirical validation. In the literature review, I realized how difficult it is to study educational interventions. I witnessed consistent issues in how studies carelessly stated conclusions when major methodological or interpretational flaws were present. By combining these two areas of reflection, I have developed an honest proposal for further research that can generate more authoritative claims.

Cognitive Dissonance

The early stage of developing this project was defining a burning question that my research was meant to answer. However, I experienced major cognitive dissonance in the early stages because I became aware of the fact that the development of a project could not actually answer any question empirically, only theoretically. This made it difficult for me to conceptualize the purpose of my project. As I began to dig into the existing literature, I became more aware of the difference between scholarship and empirical research. There is much academic writing that has been generated, some of which relies on external empirical research, but also a good portion of which relies on theoretical frameworks. As such, I began to settle my own dissonance, accepting that my project could be a scholarly product that had formidable roots in empirical research and theoretical frameworks while acknowledging its limitations for policy implications due to insufficient empirical testing of its own.

Educational Research is Unruly

During my literature review, I became aware of how difficult it is to hold educational research to high evidentiary standards. In a teacher's preparation for and engagement in teaching students, a million choices are made. This makes it extremely difficult to conduct studies of educational interventions that actually isolate the effect of the intervention. Confounding variables like unmeasured differences in non-randomized comparison groups can obscure the effect of the intervention. The measured endpoint must also be measuring something directly meaningful to the question of the study, not a weakly correlated surrogate for the effect desired. To be able to hang your hat on a result, there must be careful consideration of both confounding and endpoints.

Measured or unmeasured confounding was present in all of the studies that I examined, but at different levels of severity. Obvious initial differences in comparison groups were often disregarded, making it impossible to discern what portion of the effect was due to the intervention rather than just the differing nature of the comparison groups. For example, in the study by Blanchard et al. (2010) the pretest scores of the two comparison groups were already different, so the difference in post test scores is confounded by whatever differences were already present between the groups initially. There were also methodological issues that introduced confounding. For example, the intervention group in the study by Geier et al. (2008) received hours of professional development and ongoing support that the control group did not receive. This makes it difficult to assess what portion of the effect was attributable to the intervention alone versus how much the effect relies on the inclusion of the professional development and ongoing support. This type of confounding from methodological choices may be unavoidable in some cases, so when it is present, it requires careful wording in the conclusion. In the Geier et al. (2008) example of the additional support given to the intervention group, the conclusion must attribute the effect to the combination of the intervention and the additional support, not just to the intervention.

Additionally, endpoints sometimes lacked validation as a measure of what they were claiming to measure. If a study used a survey of student perceptions as a measure of inquiry application in the classroom, then there needed to be some empirical verification that the correlation was justified, listing an estimated value for the level of correlation. If the student perceptions themselves are the endpoint being measured, then the study results are highly informative to a research question about student perceptions. However, in the study by Liou (2021) the student perceptions from a survey are used to identify whether teachers were using an inquiry or teacher-directed approach. In that case, the extrapolation needed to be justified by some calibration showing that the student perceptions of their classroom experience actually does correlate to whether the teacher is using an inquiry approach in the classroom, perhaps verified by an inquiry expert doing classroom observations using an inquiry rubric to determine if the teacher's approach is considered inquiry or teacher-directed. Survey responses are also subjective, so they should only be cautiously used as a measure of any external metric unless the human perception itself is the defined endpoint. Unreliable or over correlated endpoints make these studies' conclusions spurious.

Finally, I began to develop an appreciation for how difficult it is to verify that the intervention being applied is actually being applied with integrity. After applying my own critical lens to the literature in my review, I propose that to make the results reliable and generalizable in an educational study I would need to have

1. A reliable endpoint that is validated to measure what is claimed to be measured. While having other limitations, standardized tests like the one used by Geier et al. (2008) do provide a more reliable signal for performance than self reporting (Bunterm et al., 2014; Liou, 2021).
2. A large enough trial for the intervention signal to get above the noise of other unmeasured covariates. The study by Geier et al. (2008) had this strength.
3. Comparison groups that have evenly distributed pre-intervention outcomes. In other words, randomization would be a valuable consideration (Blanchard et al., 2010; Bunterm et al., 2014).
4. A methodological strategy to assess the integrity of enactment of the intervention throughout the study.

5. An honest and careful articulation of its conclusion. Conclusions cannot be made outside of what the endpoint is validated to measure. Conclusions must be made with appropriate acknowledgement of confounding and methodological limitations.

With these things in mind, I have developed a roadmap of what type of research I would like to empirically validate my model as a policy-changing curricular approach.

Proposal for Further Research

The process to empirically validate the use of this model to improve educational outcomes for science students will need to involve a number of research phases. Initially, use of the model to design and conduct classroom science learning must demonstrate that it is more effective than existing teaching practices in a single teacher's classroom. Next, use of this model can be broadened to a controlled comparison study within one school. Finally, if there is sufficient prior evidence for its added value to students, use of the model in a multi-school study similar to that conducted by Geier et al. (2008).

Anecdotal Noninferiority Study

While an anecdotal noninferiority study does not account for sample bias, selection bias, reporting bias, and confounding, it is prudent to establish a proof of concept before investing time, energy, and money into a more authoritative study. The anecdotal noninferiority study would have the primary endpoint of performance on the updated Science MCA (2024 and beyond). My classroom would be the one burdened to prove noninferiority to the average Science MCA score of grade-matched Minnesota students within the same school and course. Choosing to compare to other students within the same school is an attempt to attenuate demographic confounders since one school can vary immensely from a state average. This sort of study can establish a rationale for further investment of time, energy, and money in research if the result is positive. If it is not positive, then use of the model will need to be adjusted.

Year Over Year Comparison Study

After use of the model is positively rationalized by an anecdotal noninferiority study, it is justified to run a year over year comparison study where one or multiple teachers establish a baseline of one or more years of student performance before applying the intervention, and then the following year implement the intervention and determine the added value. The endpoint remains performance on the updated Science MCA. In the year that teachers use the model, they will be required to utilize the MBI-inspired protocol and data sets at least eight times before the endpoint is assessed. Teachers will receive the same number of professional development (PD) hours in the model use year as they did in previous years, but their PD will focus on implementation of the MBI-protocol and data sets in the year that they use it. When the endpoint is assessed, the model use year must demonstrate superiority to the prior years of standard practice in Science MCA scores according to a prespecified statistical threshold. This study will improve on the initial study because it controls for many confounding variables present when comparing results of different teachers. Since it is the same teacher year over year, the impact of differences between the teacher one year and next are minimized compared to a scenario where you compare two distinct teachers. Nonetheless, due to the small number of total teachers involved in this study, it is unlikely that this study will be powered to support a policy change in science educational practice even if the results are positive due to residual unmeasured confounding and sample bias. Should the results of this study be positive, it can be justified to conduct a larger study that can provide a more authoritative conclusion on the effect size of using this curricular resource.

Large Randomized Controlled Trial

To establish a more authoritative conclusion on whether this model should influence policy for science education training and implementation, a large randomized controlled trial should be conducted. To scale the study to other policy-changing studies, the sample size could be comparable to that of Geier et al. (2008), meaning close to 5000 students, 37 teachers, and 18 schools. The endpoint will remain student performance on the updated Science MCA. Teachers should be randomized to model use or standard practice, since this reflects the possible policy recommendation of current and preservice teachers being encouraged to use this model. There should, however, be no difference in the support received by the comparison groups. They should have equal professional development time and equal amounts of ongoing mentor guidance. The model use group will be required to implement the MBI-inspired protocol and data sets 4 times before the endpoint is assessed at the end of the semester.

At this point it is particularly important to ensure fidelity to the model, so as to truly be able to ascertain its effect. The difficulty here is how to balance that methodological choice for the standard practice group. I suggest that both groups are required to have 4 class periods video recorded, where the model use group has 4 class periods of their implementation of the model video recorded, and the standard practice group has 4 of any class periods video recorded. The model use group will have their teaching scored on a rubric to assess the fidelity to the model while the standard practice group will receive school administration feedback on their teaching. A necessary overall quantitative level of integrity to the intended model use for all 4 recordings will be predetermined using the “integrity rubric.” If the implementation falls below that threshold, the data for that teacher will not be considered in the final analysis. The reason to not consider them as data points in the standard practice group is because they may be making different choices with the intention of using the model than they would without that intention, even if their use of the model does not sufficiently represent the intervention.

The final result of the Science MCA scores between the two groups must show superiority of the model use group according to a prespecified statistical threshold in order to be considered grounds for policy change. If it does not demonstrate superiority, but also does not demonstrate inferiority, then the results can be used by individuals to make their own choice of approach. If the study demonstrates that model use is inferior to standard practice using a prespecified statistical threshold, then use of the model should be reimagined with the best available evidence on science teaching.

One caution that I will note is that in the possible study designs that I have sketched, I make the assumption that the Science MCA scores are a meaningful endpoint. While it may be a politically meaningful endpoint, affecting the *perception* of success in a school as defined by the *state*, the Science MCA may not assess all of the elements that are worth caring about in educational research. Since I am not privy to the contents of the Science MCAs that will be implemented in the 2024-2025 school year in Minnesota, I will not prematurely defend their reliability or meaningfulness as an endpoint. Other metrics of worthy consideration would be critical thinking skills, attitudes towards science, self-efficacy, level of civic engagement, post-secondary life satisfaction, and post-secondary employment levels. The reason for choosing the Science MCA score in the study designs above is because of its broad saliency for schools in Minnesota over the next few years, regardless of whether it is empirically worthy of that prioritization.

Getting to seek out and appraise educational research was an important experience in my academic career. I have begun to reflect on what sort of evidentiary standard I will accept to change my own teaching practice, as well as how I would want to design my own research in the future. The experience has sparked an interest in me to pursue doctoral studies in how to hold educational research to high evidentiary standards. I have much more to learn about the breadth of methodologies used in educational research and their justifications before being able to meaningfully contribute to that topic. For the immediate future, I hope to make literature review a more consistent habit of my educational practice as I continue to teach.

My Voice in the Scholarship Around MBI

I knew initially that my project idea was going to fall under the umbrella of "inquiry" in science education. What I did not know initially was that there was a branch of inquiry that emphasized the role of representing students' thinking in the process of understanding a system. When I stumbled on research about model-based inquiry (MBI), I found exactly that. The process of developing my research foundation and then my project invited me into the scholarship around MBI, but it took a number of key experiences to find my voice in adding to the scholarship in this branch of inquiry. A Zoom conversation with an MBI trailblazer, Dr. Mark Windschitl, proved to be a challenging and thought-provoking experience. The other key experience was choosing to develop my own MBI-inspired pedagogical model instead of relying on a previous iteration presented in the existing scholarship.

Challenged by Dr. Windschitl

In my first draft of my project description in Chapter Three, I relied heavily on the learning model presented by Windschitl et al. (2008). While writing this chapter, I realized that Dr. Windschitl's email was on the front page of the article I was referencing, so I took the chance of sending an email asking if he would discuss my project with me. Surprisingly, I got a prompt response that he had time the next day. On the Zoom call, I experienced push back from Dr. Windschitl on a few aspects of my project.

The first area of dissonance between us was my use of the term "data." I generously use "data" to refer to any presentation of information, be that mathematical, model-based, or just examples and non-examples that students can mine for patterns. Dr. Windschitl held to a definition of the term "data" reserved for formal measurements or observations of a phenomenon. He was willing to see how the approach I was describing was student-centered and allowed them to practice inferential thinking from examples but saw my use of "data" to inappropriately dilute its meaning. While I appreciate the precision of his usage of the term in a scientific sense, one of the most impactful features of my educational training as a science teacher was realizing that my concept of what counted as "data" for students to wrestle with could be flexible and creative. My reflection going forward is that I believe expanding the concept of "data" for science teachers can importantly unlock educator's imaginations to teach concepts and conventions in science even when the information is not experimentally derived by the students themselves. While Dr. Windschitl and others have illuminated the pairing of modeling and experimentation in a much needed way, I do not believe that the usage of

MBI processes needs to be reserved for experimental-focused science explorations. Without the push back of Dr. Windschitl on this topic of “data,” I would not have developed the clarity with which I think about “data” in my project.

The second aspect that brought our differences to the forefront was our consideration of “phenomena” as the basis for an exploration. Dr. Windschitl held that “phenomena” are the basis of MBI and an expression of the natural world, with infinite nuances that could be represented in models with the goal of describing and explaining a system. Throughout my project, I was suggesting a model for science learning that is not always tied to “phenomena,” but rather to human conventions and models that are also important for students to explore the nature of. My use of MBI to explore science conventions and concepts that are not necessarily “phenomena” struck him as a reductionist approach to MBI since human conventions and models inherently simplify systems that have a vast amount of nuance, and thus modeling those simplified systems does not evoke the same process of deliberation for students grappling with natural phenomena. I concede to his argument here. I will not claim that examining examples of a convention or exploring existing models is equally as complex as modeling natural phenomena. Nonetheless, I will argue that using MBI processes to help students gain confidence with important non-experimental science concepts and conventions is a valuable possibility for teachers, particularly for teachers who are less comfortable in the role of *guide* and not *dispenser of information*. This challenge by Dr. Windschitl did promote me to reflect and consider how my non-experimental data sets and learning cycles could be better situated within natural phenomena contexts. His reverence for the complexity of natural phenomena inspired me to develop the *phenomena storyline* aspect of my project. I now see connecting science topics to relevant natural phenomena as an important component of developing buy-in from students and offering a greater *why?* to the explorations they are participating in. Like my discussion about the term "data," my discussion with Dr. Windschitl around the use of "phenomena" proved to be extremely valuable to me in the development of my thinking and, ultimately, my project.

Developing My Own MBI Model

In my final iteration of my project, I decided to develop my own learning model. During my project development I realized that I would be able to have much tighter cohesion within my project if I developed my own learning model rather than try to forge superficial connections to each piece of someone else's model. I also had some metacognitive reflections on the model development process that I experienced myself in the process while offering a curricular resource based in model development.

As I addressed in the previous section, my initial project description relied heavily on the learning model presented by Windschitl et al. (2008). I was planning to directly use the Four Conversations model from their article, and then incorporate an adapted list of their "artifacts of communication" to guide the teacher using the process (Windschitl, 2008). When I made the choice to create my own model to use in my project, I knew I would need to situate myself with some integrity to the previous literature on MBI. I knew I would need my model to include the key features that show up in other MBI learning cycles. In my literature review, I had already developed Table 1, where I describe the shared features of three different MBI learning cycles presented by Dass et al. (2015), Neilson et al. (2010), and Windschitl et al. (2008). I used the overarching themes from Table 1 to develop my own model. By choosing to create my own model, I had a lot more flexibility to relate the shared features with the emphasis that I wanted. I was able to center the model development process, and color code the input phases to emphasize their shared position in informing the central model development process. In the end, I was satisfied with the balance of the model being useful and internally consistent while having integrity to the shared elements present in previous MBI cycles.

My last reflection on developing my own model is regarding the metacognitive nature of the experience. My project was, by definition, the development of a model-based inquiry inspired protocol that centers the activity of model development. In a recursive fashion, my project also *needed* a model of its own to represent and relate the important features of the curricular resource. Just like my project encouraged iterative development of student models after various input phases, my own learning model presented in my project underwent various iterations in response to feedback and continued consideration. The most important learning from this metacognitive experience with modeling was a realization of how emotionally difficult it is to leave behind a model that you have put a lot of energy into and step into the uncertainty of beginning anew. When I realized I needed to develop my own model instead of using the Four Conversations model (Windschitl et al., 2008), I stalled in my project journey for a week, procrastinating the jump because of how big it felt to rework everything. I expect that this is how students will feel at times in their own modeling development process. I will carry this empathy into my teaching and will speak directly to this difficulty of the process so that students know that they are not crazy or alone in feeling overwhelmed at times.

In the end, the model that I developed consists of data sets and a MBI-inspired protocol that describes and explains the learning cycle that I believe powerfully pairs the dual goals of teaching scientific practices and teaching disciplinary content. I wanted to add to the scholarship around MBI with some integrity and believe the learning model used in my protocol lives up to that. Through my conversation with Dr. Windschitl and through developing my own learning model for the project, I discovered my voice in contribution to the scholarship on MBI. I am excited to continue to learn and contribute in this sector of science education as I further develop my own teaching practice in the years to come.

Summary

The question I centered in developing this curricular resource was this: *What kind of curricular resources effectively synthesize the parallel goals of engaging students in scientific practice and understanding disciplinary content?* To answer this question, I built on the existing scholarship around inquiry, and particularly MBI. My project is a curricular resource that pairs strategically designed data sets with an MBI-inspired protocol. This resource has immediate implications for science teachers in Minnesota while still having important limitations that do not suggest policy implications at this stage. It is my goal to empirically research the effect of teachers using this curricular resource in schools. I know that part of that journey will be to continue to research and reflect on the nature of educational research itself, particularly how to hold it to high evidentiary standards. This project by itself does, however, situate me within a larger momentum of scholarship around MBI, and meaningfully adds to the existing conversation between experts.

REFERENCES

Abrams, E., Southerland, S., & Silva, P. (2007). Introduction: Inquiry in the classroom; Identifying necessary components of a useful definition. In *Inquiry in the Classroom*. Information Age Publishing, Incorporated.

Anderson, C. W. (2008a). Perspectives on science learning. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education*. (pp. 3 - 30). Lawrence Erlbaum Associates.

Anderson, R. D. (2008b). Inquiry as an organizing theme for science curricula. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education*. (pp. 807 - 830). Lawrence Erlbaum Associates.

Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and Children, 46*(2), 26–29. https://doi.org/10.2505/3/sc08_046_02

Blanchard, M., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A., & Granger, E. M. (2010). Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education* (Salem, Mass.), *94*(4), 577–616. https://doi.org/10.1002/sce.20390

Bunterm, T., Lee, K., Ng Lan Kong, J., Srikoon, S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. *International Journal of Science Education, 36*(12), 1937–1959. https://doi.org/10.1080/09500693.2014.886347

Bybee, R. (2002). *Learning Science and the Science of Learning: Science Educators'*

Essay Collection. National Science Teachers Association.

Cartier, J. L., Smith, M. S., Stein, M. K., & Ross, D. K. (2013). *5 Practices for Orchestrating Productive Task-Based Discussions in Science*. National Council of Teachers of Mathematics.

Colburn, A. (2000). An inquiry primer. *Science Scope, 23*(6), 42–44.

Colburn, A., & Clough, M. P. (1997). Implementing the learning cycle. *Science Teacher, 64*, 30-33. https://www.jstor.org/stable/24152068

Dass, K., Head, M. L., & Rushton, G. T. (2015). Building an understanding of how model-based inquiry is implemented in the high school chemistry classroom. *Journal of Chemical Education*, *92*(8), 1306–1314. https://doi.org/10.1021/acs.jchemed.5b00191

García-Carmona, A. (2020). From inquiry-based science education to the approach based on scientific practices: A critical analysis and suggestions for science teaching. *Science & Education, 29*(2), 443–463. https://doi.org/10.1007/s11191-020-00108-8

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of*

Gray, R. & Campbell, T. (n.d.). *Model-based inquiry*. https://sites.google.com/view/modelbasedinquiry/home?authuser=0

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

https://doi.org/10.1080/00461520701263368

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1

Lee, O., & Buxton, C. A. (2010). Diversity and equity in science education: Theory, research, and practice. New York: Teachers College Press.

Liou, P. (2021). Students' attitudes toward science and science achievement: An analysis of the differential effects of science instructional practices. *Journal of Research in Science Teaching*, *58*(3), 310–334. https://doi.org/10.1002/tea.21643

Lonergan, R., Cumming, T. M., & O'Neill, S. C. (2019). Delivering inquiry learning in science classrooms: A planning tool. *Teaching Science (Deakin West, A.C.T.)*, *65*(2), 14–22.

Millar, R., & Driver, R. (1987). Beyond processes. *Studies in Science Education, 14*, 33-62. *Research in Science Teaching*, *45*(8), 922–939. https://doi.org/10.1002/tea.20248

Minnesota Department of Education (MDE). (n.d.a) *Academic Standards (K-12): Science*. Retrieved April 15th, 2023 from https://education.mn.gov/MDE/dse/stds/sci/

Minnesota Department of Education (MDE). (n.d.b). *Graduation Requirements*. Retrieved April 15th, 2023 from https://education.mn.gov/mde/dse/gradreq/

Minnesota Department of Education. (2010). *Minnesota Academic Standards: Science K-12: 2009 version*. Minnesota Department of Education. Retrieved April 26, 2023 from https://education.mn.gov/mdeprod/idcplg?IdcService=GET_FILE&dDocName=005263&RevisionSelectionMethod=latestReleased&Rendition=primary

Minnesota Department of Education. (2021). *Minnesota Education Statistics Summary* (2020 - 2021). [Data set]. Minnesota Department of Education. Retrieved April 15, 2023 from https://public.education.mn.gov/MDEAnalytics/DataTopic.jsp?TOPICID=4

National Research Council. (1996). *National Science Education Standards : observe, interact, change, learn*. National Academy Press.

National Research Council (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. National Academy Press. https://doi.org/10.17226/9596

National Research Council. (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas.* National Academies Press.

NGSS Lead States. (2013). *Next generation science standards: for states, by states.* National Academies Press.

Neilson, D., Campbell, T., & Allred, B. (2010). Model-based inquiry. *The Science Teacher (National Science Teachers Association)*, *77*(8), 38–43.

Ogan-Bekiroğlu, F., & Arslan, A. (2014). Examination of the effects of model-based inquiry on students' outcomes: scientific process skills and conceptual knowledge. *Procedia - Social and Behavioral Sciences*, *141*, 1187–1191. https://doi.org/10.1016/j.sbspro.2014.05.202

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science:

creating connections. *School Science and Mathematics*, *109*(7), 394–402. https://doi.org/10.1111/j.1949-8594.2009.tb17870.x

Rahman, T., & Lewis, S. E. (2020). Evaluating the evidence base for evidence‐based instructional practices in chemistry through meta‐analysis. *Journal of Research in Science Teaching*, *57*(5), 765–793. https://doi.org/10.1002/tea.21610

Ross, D., & Davidson, A. (2020). Integrating the five practices and model-based inquiry to facilitate meaningful science discourse. *Science Scope (Washington, D.C.)*, *44*(1), 43–49.

Soysal, Y. (2020). Determining the mechanics of classroom discourse in Vygotskian sense: Teacher discursive moves reconsidered. *Research in Science Education (Australasian Science Education Research Association), 50*(4), 1639–1663. https://doi.org/10.1007/s11165-018-9747-2

Vorholzer, A., & von Aufschnaiter, C. (2019). Guidance in inquiry-based instruction - an attempt to disentangle a manifold construct. *International Journal of Science Education*, *41*(11), 1562–1577. https://doi.org/10.1080/09500693.2019.1616124

Vygotsky, L. & Cole, M. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Wiggins, G. P., & McTighe, J. (2011). *The understanding by design guide to creating high-quality units.* Pearson.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education 92* (5): 941–967.

Zuckerman, G. A., Chudinova, E. V., & Khavkin, E. E. (1998). Inquiry as a pivotal element of knowledge acquisition within the Vygotskian paradigm: building a science curriculum for the elementary school. *Cognition and Instruction, 16*(2), 201–233. https://doi.org/10.1207/s1532690xci1602_3